APPENDIX 9: PACIFIC COAST GROUNDFISH FMP HABITAT USE DATABASE USER MANUAL FOR VERSION 15B (DRAFT)

Table of Contents

1 DATABASE PURPOSE

The Pacific Habitat Use Relational Database has been developed to provide a flexible, logical structure within which information on the uses of habitats by species and life stages in the west coast groundfish species complex can be stored, summarized and analyzed as necessary. This will form an important component of the information base for developing the EIS for the Essential Fish Habitat amendment to the Pacific coast groundfish fishery management plan.

The database is designed primarily to capture the important pieces of information on habitat use by species in the Pacific Groundfish FMP as contained in the Updated Life History Descriptions document compiled by NMFS. This document contains information on each of the species in the groundfish FMP that includes range, fishery, habitat, migrations and movements, reproduction, growth and development and trophic interactions. Certain elements of this information need to be captured in a database so that habitat use data can be analyzed both by species and habitat to provide input into various components of the analysis of EFH, HAPCs and fishing impacts.

Appendix 8A contains an extract from the Updated Life History Descriptions document for canary rockfish (*Sebastes pinniger*). Parts of the text in this extract have been highlighted as an example of the types of information that need to be entered into the database.

Appendix 8B contains a list of tables and forms used in the database.

2 DATA STRUCTURE

It is essential for users to grasp the principle of data structuring and how it is used in a system like this to both enforce data quality and form the basis for developing interrelated lines of analysis. It is a different concept from a simple file storage system that can only receive, store and regurgitate data for use elsewhere. This system can of course be used in that way as well but that is only utilizing a fraction of its capabilities. Appendix 8C explains in detail these essential basic principles that underlie the design and construction of this Habitat Use Database.

Figure 10 is the 'Entity Attribute Relationship' analysis diagram for the database. It shows the data tables, their fields and which of these form the 'primary keys' (in bold) and the foreign keys which link the tables together via a network of one-to-many relationships. The tables contain all the data in the database. Some contain primary data (e.g. SpeciesLifeStage and PlaceTime) and others contain reference information such as Species, which is simply a list of all the species in the FMP. All data entry forms, data checking procedures, and queries are based around this table structure.

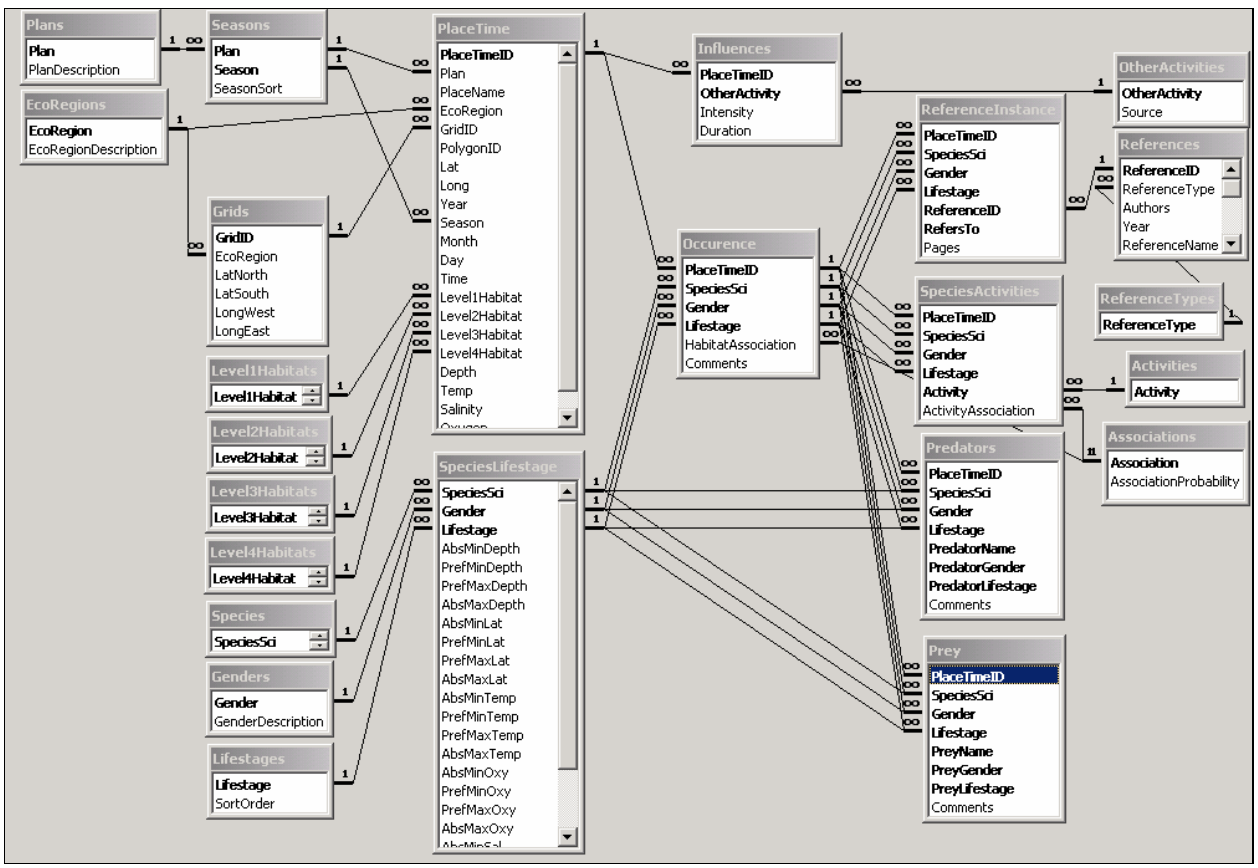


Figure 1: The structure of the data tables, their constituent fields and relationships between them.

2.1 Spatial and Temporal Data: The PlaceTime Table

The core of the database is the 'PlaceTime' table. This records where and when particular observations of species-habitat associations were recorded. Records in this table represent the place (or area) and time (or period) of the recorded occurrence of the species and life stage data and the habitat and physical conditions that prevailed at that time and place.

The principle is that the data being recorded are associated with some sort of time and space framework, whether this is in the most general sense such as the whole West Coast region for all time down to very fine scale data where exact times and places are known and might be used to stratify analyses. The system is therefore not dependant on exact spatial and temporal information about a particular species-habitat association. It can be used even where there are no spatial or temporal elements in the data. The information in the PlaceTime table simply allows the breakdown of analysis of species-habitat associations on a finer spatial and / or time scale than the entire range of the species/life stage should the information be available at that resolution. More detailed explanation of the implications of the different grades of spatial and temporal data are given in the following sections.

To allow such flexibility in the type of time and place data that can be recorded and to allow the combination of different types in the same table and analyses, it is necessary to uniquely identify each record in the PlaceTime table, referred to as 'PlaceTime' record in the preceding, with a unique number 'PlaceTimeID.' This forms the primary key in the table and cannot be repeated. **This means that either data should only be entered in one place or if there are multiple data entry sites then they should either co-ordinate with one another to ensure they use unique sets of numbers or access a centralized database via a network (local or wide area) or via the internet using active server pages. The other possibility is for the database to be 'replicated' and later 'synchronized.'**

The remainder of the fields in this table can either by typed in directly or selected from the combo boxes provided at either table or form levels. There are also range limits on temp, salinity, depth, oxygen, latitude and longitude when their values are not null.

Frequently there is no temporal or spatial information and there may be just a series of observations of species occurring on different habitat types. We don't know when or where these observations were taken, only that they are accurate in their recording of the types of habitat on which the species were seen. In such cases the record has an arbitrary but unique identifier in its PlaceTimeID field which has nothing to do with place or time but simply allows the habitat data to be linked to the occurrences of species and their activities (tables 'Occurrences' and 'SpeciesActivities').

Obviously for any of the given 'PlaceTime' records (even if it had very detailed location and date-time data) there can be multiple occurrences of different species and different life stages of the same species. These can also have multiple species and life stages of both predators and prey. The database is structured in such a way to allow the correct representation of such natural one-to-many relationships between entities.

In the PlaceTime table, the column PlaceName allows the use of place names where these are used to identify a known area or location at which observations have been made of species/habitat associations. Provided these names are used consistently (a reference set could be defined in a ‘look-up’ table) then they could also be used in a stratified analysis. This can be used independently, or in conjunction with grids and “EcoRegions.” EcoRegions are used as a simple large scale subdivision of the area covered by the FMP so that analysis of habitat use can be broken down at a finer scale than the entire Pacific coast. Seven EcoRegions (numbered 1 to 7) have been proposed, as illustrated in Figure 11. EcoRegions are defined by their member GridIDs. In this implementation of the database no GridIDs have been identified, so EcoRegion and GridID are the same (i.e. there is only one grid per EcoRegion). Arranging it in this way means that if in the future Grids are defined, there will be no need to alter any code in queried that use the Grid/EcoRegion structure. These will run without modification both with the present scheme and when the grids are reassigned.

As shown in the data model, the allocation of results to Eco-Regions should, preferably, always be done via the Grids table. This allows the flexible re-definition of eco-regions and the grid squares they contain should this ever be necessary. There is also an EcoRegion field in the PlaceTime table into which the user can enter the value of the eco-region directly and simply analyze via this field when ignoring the Grid system. There is also a PolygonID field available in the ‘PlaceTime’ table for recording finer scale spatial allocations, should these be required.

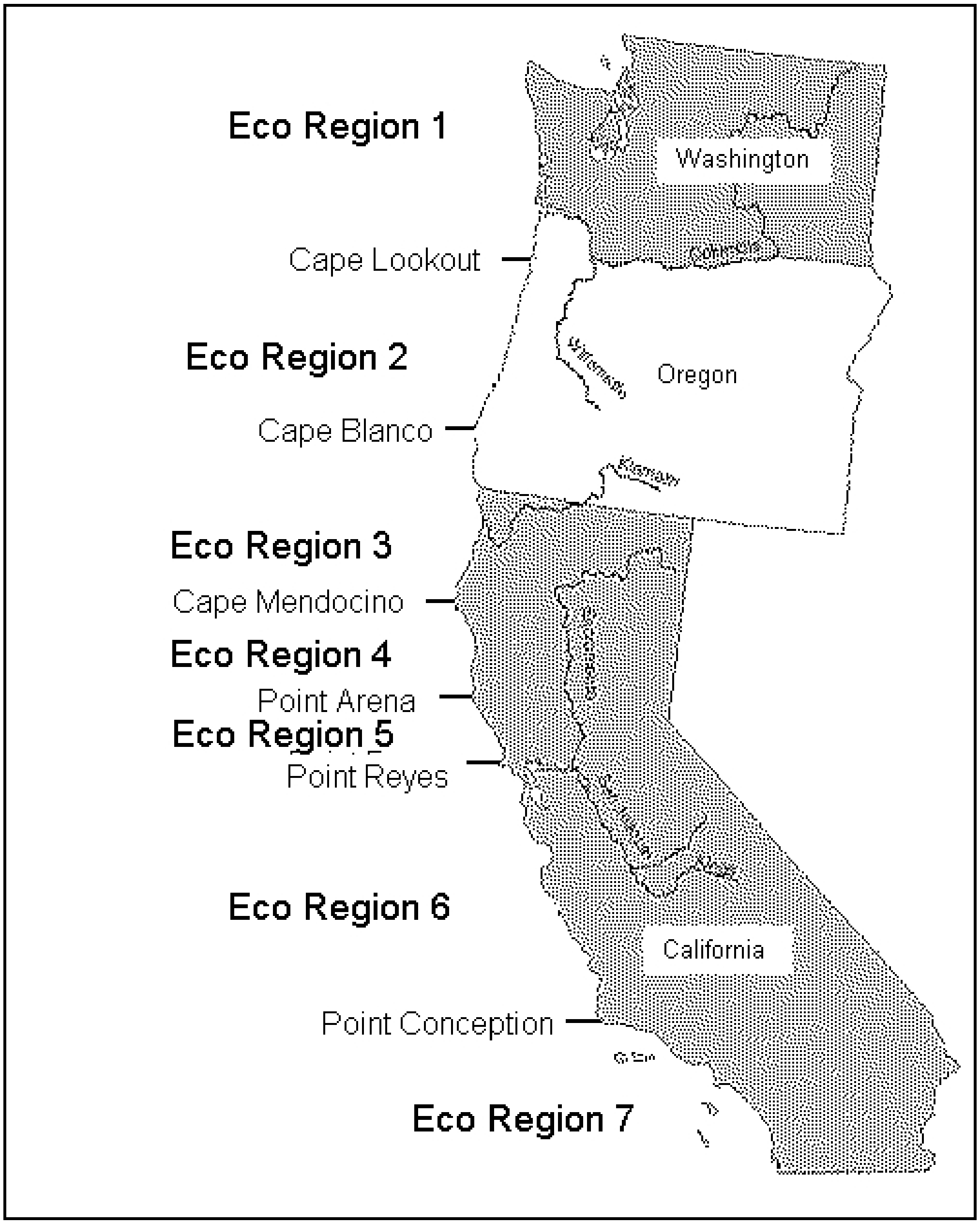


Figure 2: The Eco-Regions

The Grids table has four fields LatNorth, LatSouth, LongWest and LongEast which can be used to define a grid square in terms of its binding latitudes and longitudes. It is assumed these will be entered in a decimal as opposed to sexagesimal notation and that the upper bound for one limit will not run into / overlap with the lower bound of the adjoining limit. The GIS conventions will define the appropriate usage. As with the PlaceTime position fields, there are range limits on what can be entered based on the latitudes and longitudes that enclose the whole west Coast region.

Temporal data in the PlaceTime table include years, seasons, months, days and exact times. Data on these attributes can be entered as and when they are available and deemed to be relevant. The fields can be ignored when the data are unavailable or deemed to be irrelevant. The availability of the fields within the system and the way they are defined allows flexibility in this respect.

2.2 Scaling of spatial and temporal data

The hierarchy of detail available in the PlaceTime table allows data of different temporal and spatial scales to be combined in the analyses. It is important to bear in mind that a few basic assumptions must be adhered to in order to make informed use of this flexibility.

1. Data should be unique. Where data are collected on the basis of one of the finer temporal and/or spatial scales and are also available as a summary of this on one of the higher scales then the data should be entered into the database according to only one of these scales and preferably the finest scale available.
2. Where there are data of mixed temporal and/or spatial scales then care must be taken in framing analyses on two counts:
   a) when such data are combined in an analysis then the results can be stratified spatially or temporally down only to the level of the data with the broadest spatial and temporal scales, and
   b) when a stratification of results is intended on a fine scale, then 1) either all the data should have values entered for those fine scales or 2) careful conditions need to be set to exclude records that do not have values for those finer scales. Note, however, that in this latter case the analysis would not be using all of the available data.

2.2.1.1 Seasons

Seasons are defined within the management plan though it is not obligatory to utilize either of these features where they are not required or are irrelevant. It allows several concurrent seasonal regimes to be defined where management plans are based around a major species and the recognized seasonal patterns of these are different even though they occupy the same areas and times. Equally the defined seasonal regimes for different plans can also be matching, which is the simpler and more likely scenario. Where there is either no defined management plan or a single management plan, the structure allows the simple definition of a single seasonal regime. Where there is no information on seasons, or seasonal attributes are not applicable or irrelevant then the user can enter an appropriate single ‘seasonal’ value in the look up table such as ‘All Year’ or ‘Unknown’ or ‘Not Applicable’ or whatever the user chooses.

Should it ever be required in the future to extract or ‘manufacture’ the spatial and temporal data from the descriptive information in source documents then an example methodology is provided in Appendix 8D.

2.2.1.2 Fishery Management Plan

The system is designed to be able to represent several Fishery Management Plans by specifying the FMP in the filed “Plan” in the PlaceTime database table. The facility thus offers the

opportunity to stratify analyses according to FMPs where this is required. The present implementation does not require such a facility (there is only one FMP) but it has been left in the database structure in case data from another FMP are entered into the same database at some future time. Its functionality can be ignored by always entering one single value for the 'plan' field. **As with all such 'look-up' data values (e.g. species names etc), if the names are altered the alterations are automatically 'propagated' throughout the entire database doing away with the need to manually update all the associated data with any such name changes.**

2.2.1.3 Habitats

Habitat is currently defined in the PlaceTime table under four tiers of classification. The four levels of habitat classification are currently independent and are not structured as sub sets within one another. For ease of data entry and comparison all three levels are displayed within the same form (Figure 12). As with all of this kind of 'look-up' data the user is free to add or alter the values under these classification schemes.

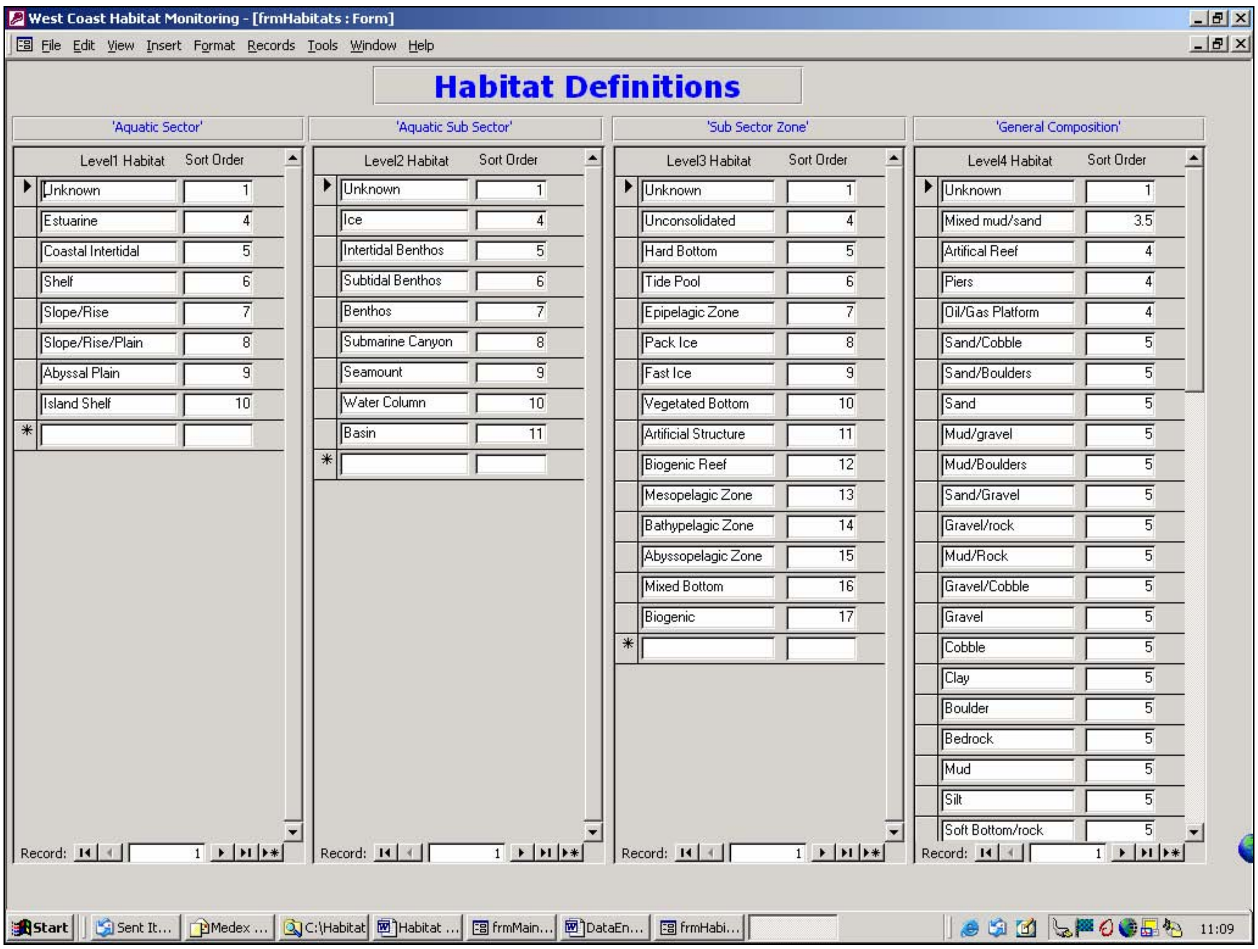


Figure 3: Habitat definitions

2.3 The Species, Genders and Life Stages Tables

These data each reside in a separate table and also in a combined SpeciesLifestage table, although the forms that serve these tables have been conveniently combined. There is also a button to call up the life stages form so that the life stages available in the Lifestages table can be added to or amended. The same is the case for genders. The design of the system also assumes that all predator and prey species and life stages are also entered in these tables. Where no life stage info is available or is deemed irrelevant then a value such as 'Unknown' or 'All' must first be entered via the Lifestages form. This will then appear as a life stage option when entering the value of the life stage for that species under the SpeciesLifestage form. The same principle applies for genders. Care must be taken to bear both data entry and later analyses in mind when deciding on values for life stages and genders at the data entry stage and on what values to filter these on during analyses, e.g., if a combination of 'Both' and 'Unknown' are used as values for gender then one or the other must be used alone with reference to a particular life stage and not both of them. If you used both of the two values it could conceivably distort results. Equally where 'Both' and 'Unknown' have been correctly applied as genders to different life stages then the two values must be used in any filter that is being applied across genders and life stages for a given species.

2.4 The Occurrence Table

The 'Occurrence' table records which species and life stages occurred in the recorded place and time frame on the recorded habitat, etc. The relational structure allows the recording of several of the life stages of the same species that may occur simultaneously and of course as many species as there were present. As explained earlier if no spatial or temporal data are available then the so-called PlaceTimeID simply refers to the habitat type only, as defined.

There is also a 'HabitatAssociation' field in the Occurrence table which records a measure of the relative strength association of that species-life stage with the habitat recorded (as strong, medium, or weak) with matching probabilities. The number and names of values and the probability figures can be changed by editing the Associations table via its form or directly in the table. All the values are the same as those presented for the degree of association of a particular 'Activity' as well.

2.5 The SpeciesActivities table

The SpeciesActivities table records the activities of the fish (spawning, breeding, feeding, or growth to maturity) on a particular habitat in a particular time and place. There may be multiple activities for any given species-life stage in a Place/Time frame. As with the habitat associations, the degree of association of that activity performed by the fish in that habitat can be recorded as strong, medium or weak.

Associations between species can be derived via a query that groups which species-life stage-activities were occurring in a given Time and Place frame for the various habitats. This is providing all data have been comprehensively entered.

2.6 The Predators and Prey Tables

The predator and prey tables have a many to one relationship with the Occurrences table. i.e. any one given species at a particular life stage can have many predators and can also itself prey on several other species. These predators and prey will themselves also be at a particular life stage. The predators and prey recorded must also be represented in the three tables ‘Species’, ‘Lifestages’ and ‘SpeciesLifestage’ even if they are not in the FMP species list. For convenience and simplicity of design the main predator and prey groupings have been denoted as either a member of a predator (pred) or prey grouping in the comments field. That field is then sorted in the menu choice so that these groupings appear together.

2.7 The Influences and OtherActivities Tables

The database also accommodates the recording of other activities or occurrences (impacts) that might have influenced species and their activities in a particular time and place. This is done through a sub-section ‘Influences’ on the bottom of the ‘Place-Time Centric’ by allocating these “OtherActivities” in the “Influences” table the same PlaceTimeID as in the PlaceTime table. The extent to which this facility will be used is not clear at present, but this structure will allow comparative analyses to include such influences or ‘impacts’ as well as habitats and the other attributes on patterns of occurrence and species activities at their various life stages.

Such things as Pelagic Fishing or Acoustic Surveying can be recorded but also natural events such as an el-Nino event or a turbidity current. The OtherActivities table also has a field ‘Source’ that allows the user to group these other activities according to their source. This can be employed flexibly as required. E.g. it could take only two values such as ‘Human’ and ‘Natural’ or these could be subdivided further as required according to the kind of analysis being undertaken. As with the occurrences table the value of the PlaceTimeID is automatically inherited from the parent PlaceTime table in the form used to enter data.

2.8 References

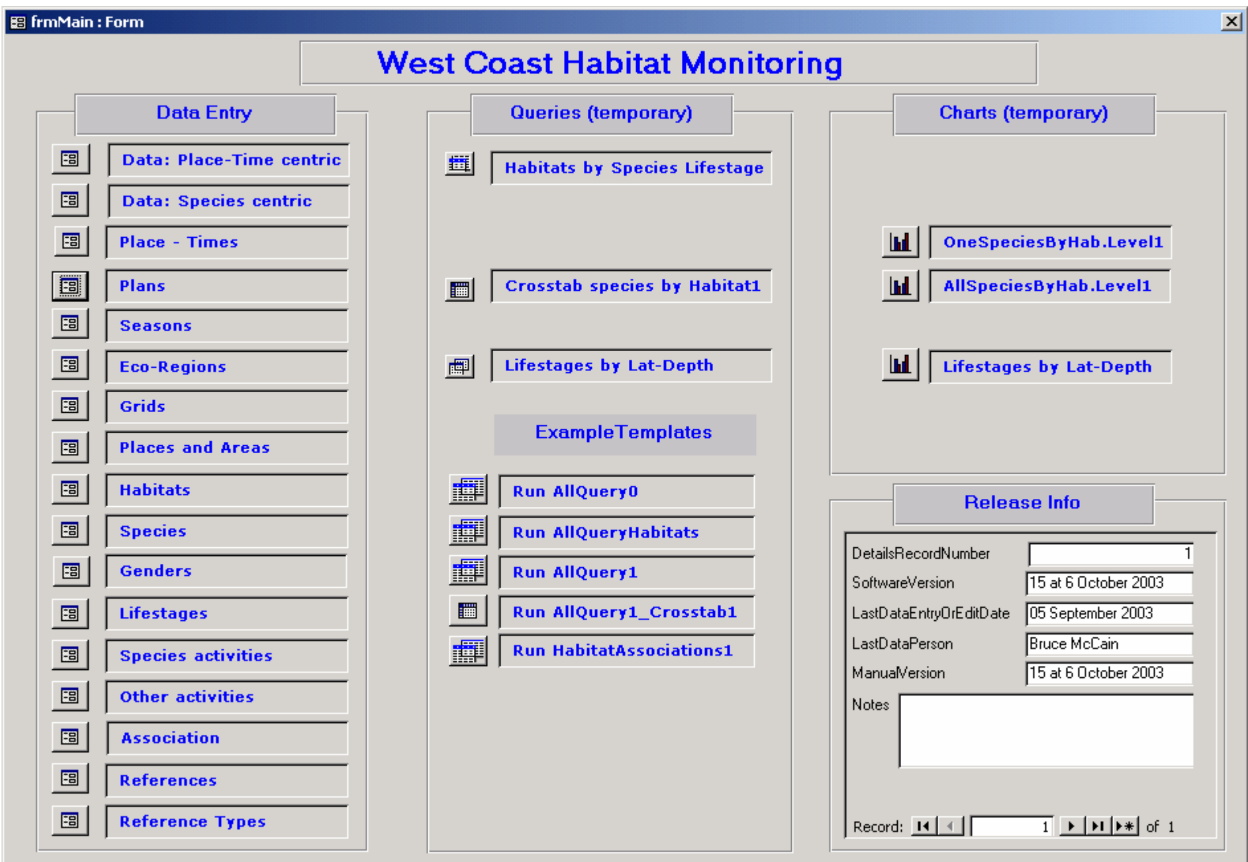
All reference materials are recorded in a single table “References.” Each work should be recorded only once with a unique identifier ‘ReferenceID’. The ‘ReferenceInstance’ table records the occurrences of that reference as and when it is referred to in relation to a given occurrence of a species-life stage for a specific time-place frame with its associated habitat and physical conditions. Thus a given reference can appear as many times as necessary in the ‘ReferenceInstance’ table even for the identical PlaceTime, Species and Lifestage providing it refers to different aspects as recorded in the remaining key field ‘RefersTo.’ For example, the same work can be recorded as a relevant reference for both Habitat and Predators.

A total of 557 references have been entered so far (October 2003). These are then also referred to from the database, thus explicitly describing the network of references and the context in which they are referred to.

3 WORKING WITH THE DATABASE

The database is designed to be as intuitive as possible with information naturally arranged in a hierarchy of 'Parent' – 'Child' tables. These tables are automatically linked in their data entry and viewing forms. For those unfamiliar with Access databases a period of practice on a dummy copy of the system will help familiarize the user with navigational controls. Liberal use of the 'Help' button should also be made.

The opening form appears as:



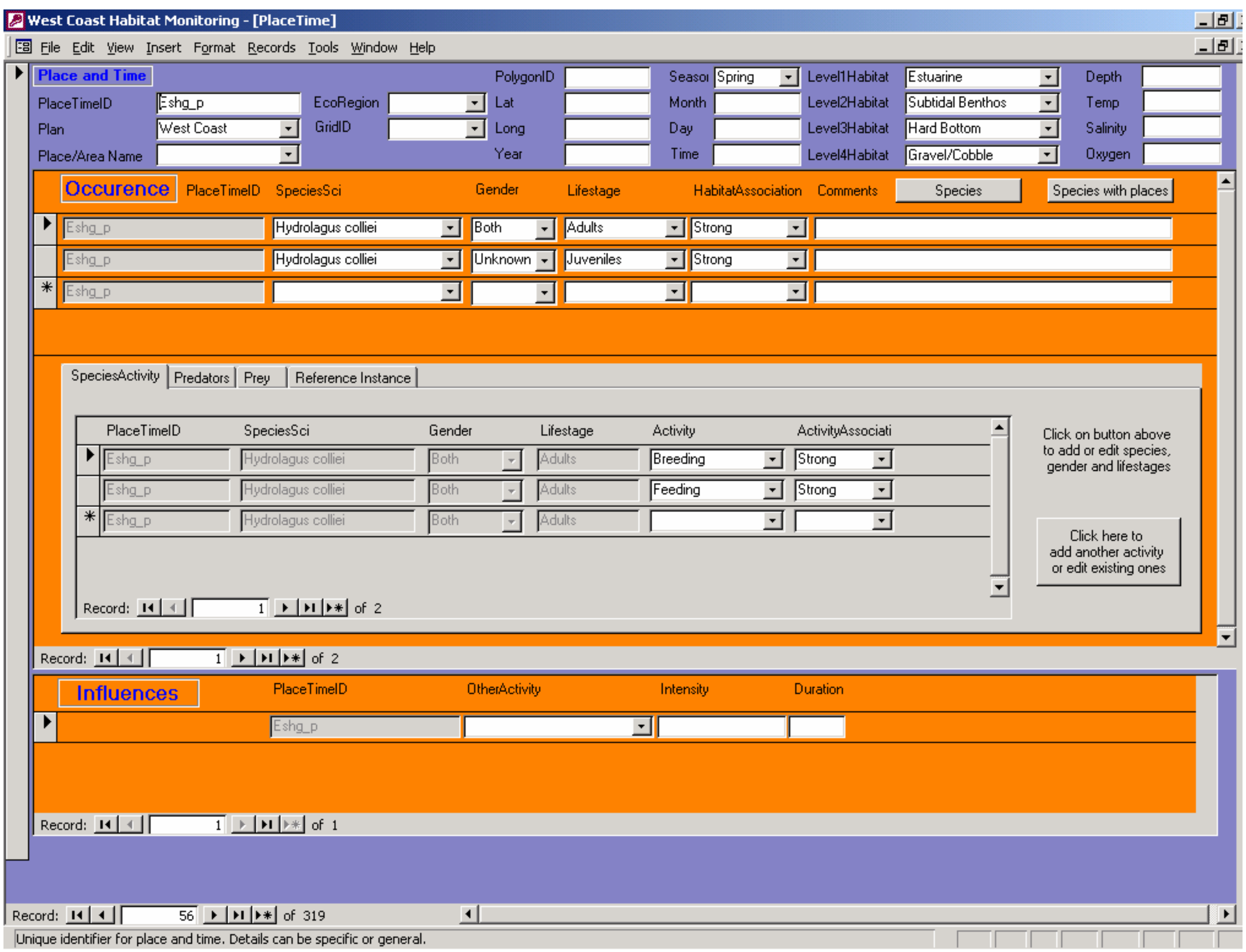
This form lists all the current options for data entry and data analysis. Additional queries and charts can be developed as required.

The 'Release Info' section presents a summary of which version of the software is under use and which data set it incorporates. This aims to reduce the danger of any copies getting out of synchrony with one another where data entry and analysis is ongoing at a number of sites. It also helps ensure the users have the correct set of documentation to go with the product.

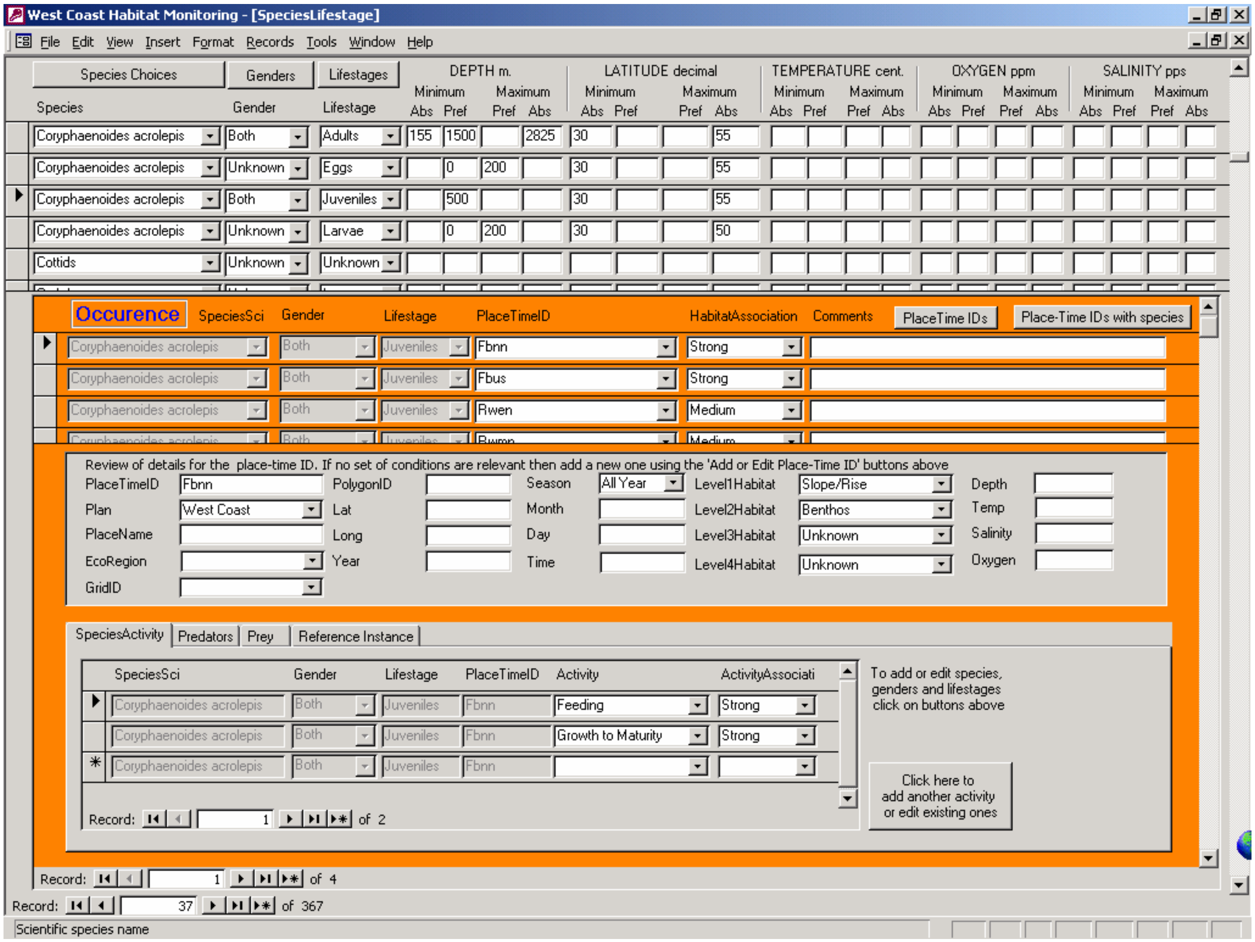
3.1 Data Entry

There are two main options for data entry: "place-time-centric" and "species-centric." Essentially the first allows you to enter all the species-life stages for a given habitat whereas the latter allows you to do the converse and enter all the habitats for a given species-life stage. The 'Place-Time' scenario is more likely with data arising from a survey while the 'Species-Centric' approach is more likely with data arising from a literature survey. In both instances the associated, more detailed, place and time info can also be entered to the degree in which it is available. In both cases, data entry starts with a Main Data Entry Form (see below). This form is arranged in sequential sections, as emphasized by the different colors. It is important to use the correct set of record navigation buttons for each section. In the Place-Time version, the top level records for 'Place and Time' have a blue background with their record navigation buttons at the bottom of the form; note that there are four sets of navigation buttons in this form. The next two sections, Occurrences and Influences, are nested at the same (2^{nd}) 'level' and have a copper colored background. Nested within Occurrence at the third level are four sub tables, each with an independent serving form, Species Activities, Predators, Prey and Instances of References.

The place-time-centric main data entry form appears as follows:



The species-centric main data entry form appears as follows:



The species version shows similar information, but it allows data entry by species and life stage, to simplify the process of entering data from the Updated Life History Descriptions document, which will be the primary data source in the first instance.

Whichever of the forms is used, the data always end up in the same underlying data tables in a unified and consistent data structure. The only difference is how this is shown in the user interface.

In both cases, all of the various sections of the form are synchronized. Thus when the user moves on to the next place-time or species record all of the associated data that have already been entered automatically appear in other parts of the form. Note that if new data are being entered then the correct matching PlaceTimeID and Species_Sci/Life Stage are automatically copied to the occurrences table. This principle also applies to all the other linked tables at the lower levels with their key fields.

Multiple ‘Occurrences’ and ‘Influences’ can be visible for any one PlaceTime record. These two logically occupy the same level in relation to the parent PlaceTime record and are given matching background colors in the Place-Time centric to visualize that fact.

Within the ‘Occurrences’ of a single species-life stage there can be multiple activities that the species-life stage is performing on the recorded habitat. There can also be multiple predators and prey in that location and multiple references relevant to that occurrence. All of these data elements are recorded on the tabbed sub forms. These have an ‘index-card’ like appearance to maximize the amount of data available on one screen.

Most of the data are presented in forms in a table-like format with multiple rows for records. This is considered to be the most useful and practical approach for the user who is entering and reviewing data. Often corrections (and also avoidance of typing errors) involve the comparison of adjoining records, especially when the data in question has been filtered and sorted. The table-like interface is far more useful for doing this since everything is visible at once.

Having all of the inter-linked data from related tables visible at once in adjoining sections also prevents confusion and errors during data entry and simplifies the making of corrections and/or modifications after the data have been entered. It is impossible to enter the wrong data in the key fields for related tables since the foreign key constraints automatically generate an error message when the user tries to do this. The form arrangement in any case does away with the need to re-type related key field data since it is automatically copied from the ‘parent table’ section to the ‘child table’ section of the form and cannot be edited there but only viewed.

Appendix 8E is a ‘tutorial’ explaining how the information for a given species is broken down and entered as records.

The database system has its own tool bar:

Under the West Coast heading the entire functionality of the main control form is reproduced so that users can call up any data entry form or analysis direct from the menu bar without having to re-locate the opening form. All of the important functions on the tool bar and many others are also provided by the main menu bar. It is therefore not essential to use this ‘WestCoastTools’ toolbar to operate the system and it can be turned off under the menu choices ‘Tools/Customize/Toolbars/WestCoastTools’ should the user prefer not to use it. The tool-bar can also be turned on and off by right clicking on the empty area to the right of the main menu bar at the top of the screen and then ticking ‘WestCoastTools’ on or off.

The user can unhide the main database window in order to access the underlying parts of the system directly. NB **Changes at this level should be made only by an experienced database designer or code developer who is responsible for the database.** This is especially important if there are multiple copies of the database being used which need to be synchronized. This could be where data are being entered at several different sites or data entry going on at one sight and query development at another. In such situations, requests for alterations and additions to the

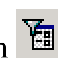
system should be first logged and then implemented on an organized basis so system development and the data management can proceed in a consistent and integrated fashion across sites.

3.2 <u>Look-Up Tables</u>

Look-Up data are those provided in tables such as Species, Genders, Lifestages, Eco-regions and Grids, Habitat Levels, and Activities. These data change less often than those in the other tables. New records are only occasionally added and existing ones are only rarely altered. When changes are made these immediately become available as data entry choices in the main data entry forms. When they are altered, all the records in the database that have the old values are updated automatically to reflect the change. Note that you can not delete one of these look-up values unless you have first deleted any records elsewhere in the database that refer back to it.

For convenience, some of the more likely look-up tables can also be called up from the main data entry forms (PlaceTime and SpeciesLifestage) via various buttons and also as sub-choices under the WestCoastTools tool-bar.

3.3 <u>Sorting and filtering data</u>

One important aspect is learning how to use the sorting and filtering buttons. A user can filter the data so that only records appear that have field values equal to that of the field the user is currently in. This is known as ‘Filter by Selection’ . Secondly a user can filter by form  by first selecting this button and then choosing from the list of available values provided by the drop down boxes that become available for ALL fields. The user would then press the apply filter button  to obtain the subset of data. The term ‘(filtered)’ appears next to the record counters at the base of a form / table whenever a filter is in operation. Remember to check this and clear the filter afterwards by pressing the same  button again. A user can also remove the filter completely with the  ‘clear-filter’ button. In the filter design view a user can also clear the filter grid. Most of the data are already sorted according to its key fields. In some instances there are additional sort-order fields e.g. for life stages or seasons. This allows the order that the values appear in to the user to be assigned even when using normal descriptive terms. The user can resort the data according as desired as an aid to locating particular records during editing etc.

3.4 <u>Analyses</u>

3.4.1.1 Overview

These are currently under development. It has been requested to provide only a few working examples of the different types of query with documentation of how these are developed and can be adapted and extended. The client then intends to use these as the basis for developing their own queries.

Attention is drawn to the sections on Data Structure and Appendix 8C which detail the essential principles and knowledge required to make the best use of this system's capabilities in this respect.

Examples of a select query, a cross-tab query and a chart which plots the results of a cross-tab are provided. Other analyses that provide lists of species and life stages according to the various habitat categories, grid squares and eco-regions can be developed if the data are broken down to this extent in the future.

The queries in the examples are plotted via a generic method whereby axes labels are formed from the category values themselves and thus always reflect the data content. Thus there is no need to create a separate explicitly labeled chart each time the selection conditions or underlying data change.

Complex patterns of trophic interdependence are represented via the conjunction of the Occurrences, Predators and Prey tables. With some thought it may be possible to develop queries that can analyze these patterns.

Where analytical requirements demand the use of mathematical and statistical modeling software, queries can be developed to produce the correctly formatted data-sets for direct input into such applications. An example of this is provided with the **'HabitatAssociations1'** query.

Further queries could also be developed to interface the database with companion systems which could both receive and provide data in integrated analyses.

If data are provided there is the opportunity to analyze for the recorded 'Influences' (or 'impacts') where these may be natural or anthropogenic.

To concentrate on the scope of the data provided so far, a series of examples follows covering different classes of queries with explanations of how these were developed and how they can be extended. In addition to the detailed instructions given here it is recommended that anyone developing such queries should have a clear grasp of the principles of relational databases and query structuring and have good Access manuals or text books available. Beware that many of the text books place the 'cart before the horse' and embark on detailed 3rd generation code examples without first clearly explaining the essential underlying relational principles of such 4th generation database environments. Despite such systems being available for 20 years or more, by and large within biological resource management the penny has still not dropped! A very good reference work would be 'Access Database design and programming' by Steven Roman, published by O'Reilly, ISBN 1-56592-626-9.

A user should not alter any of these example queries, but should instead copy and rename them and then experiment on those new queries using them as a template to develop new lines of query. That way if it all goes horribly wrong the user can simply delete them go back to the unaltered source queries and copy them afresh. In addition, a strict and documented system of regular backups should be in place as well.

A user can copy and rename the queries by right clicking on them in the queries section of the main database window, selecting 'copy,' then right clicking in an empty space in the database window and selecting copy, entering a new query name, and then working with that new query. One important thing to bear in mind is that if a query uses other queries as its source and you rename those sources then obviously the name of the source also has to be altered in the query that uses it. E.g. the query called 'AllQuery1_Crosstab1' which uses 'AllQuery1' as a source. You can also 'import' individual queries from backups if you inadvertently damage one of these source queries.

Some charted output is provided as part of the system. This is mainly to demonstrate its capabilities in that any output can be charted where appropriate. It is beyond the scope of the resources available for this manual to explain in detail how to develop charts. Also, developing the kind of generic charts demonstrated here, that label their own axes etc according to dynamic inputs, assumes some expertise in the use of Access. Any additional charts required could be developed and provided in future.

3.4.1.2 Example 1: Species-based investigation

In this example we will develop a range of queries that will look at all the available data for a particular species. Obviously one of the conditions will be the species name. Thus the queries can simply be reapplied to any other species by altering the value of the species name under that condition. The main query will stratify according to all life stages. Genders will be ignored in this case as there is very little gender specific data that has been entered so fare. For each of the life stages we will look at each defined habitat in turn and list its definitive values. Within those 'strata' we will then analyze for species activity, predators, prey and even references in the literature. In summary the complete list of strata are:

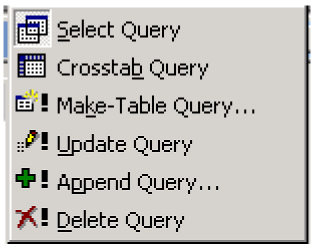
- Species
  - Life stage
    - Habitat
      - Species Activity
      - Predators
      - Prey
      - References

Note that the last four are all 'independent' attributes within habitat.

The example query is called 'AllQuery0.' The query is created from the 'Queries' section of the main database window and selecting 'Design View.' One can make good use of the 'Simple Query Wizard' and 'Crosstab Query Wizard' provided one has sufficient database experience. Care is required because it is possible to produce a working query that provides results that are nonsense if tables are linked and conditions combined in an illogical fashion.

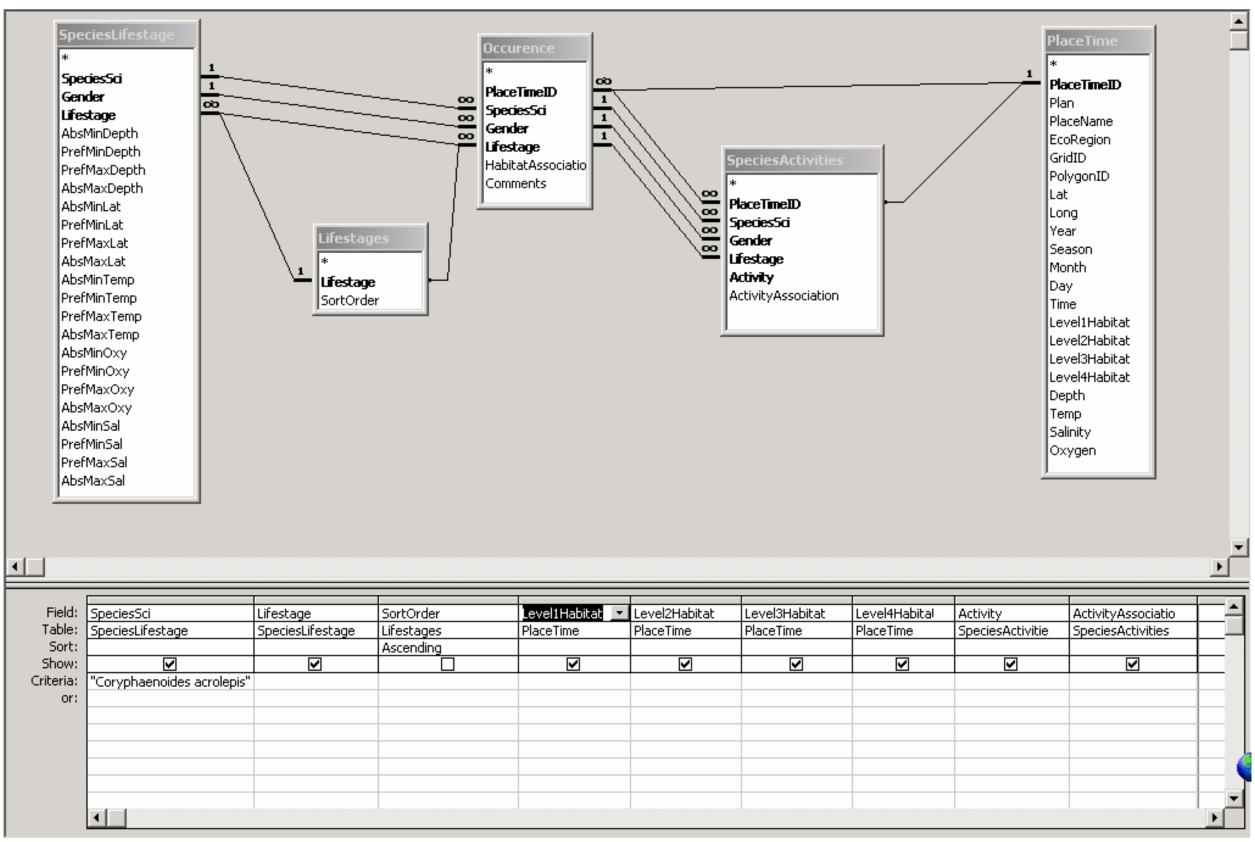
Tables are added to the design view by selecting them from the 'Show Table' list offered. Note that a user can also base a query upon another query as well.

Apart from the selection and cross-tab queries that produce a set of results a user can also alter the query type so that it adds, modifies or deletes data or even creates new tables. Make sure you know what you are doing before going down that road and have in place a religious back up procedure that is fully documented so that you can reverse out of any inadvertent disasters!



You should select the tables you require for your query by first looking at the 'map' of your database provided by the 'relationships' diagram.  This will serve to remind you of table and field names and how they relate to one another.

You should select the tables for the query from the 'Show Table' to produce a query as follows.

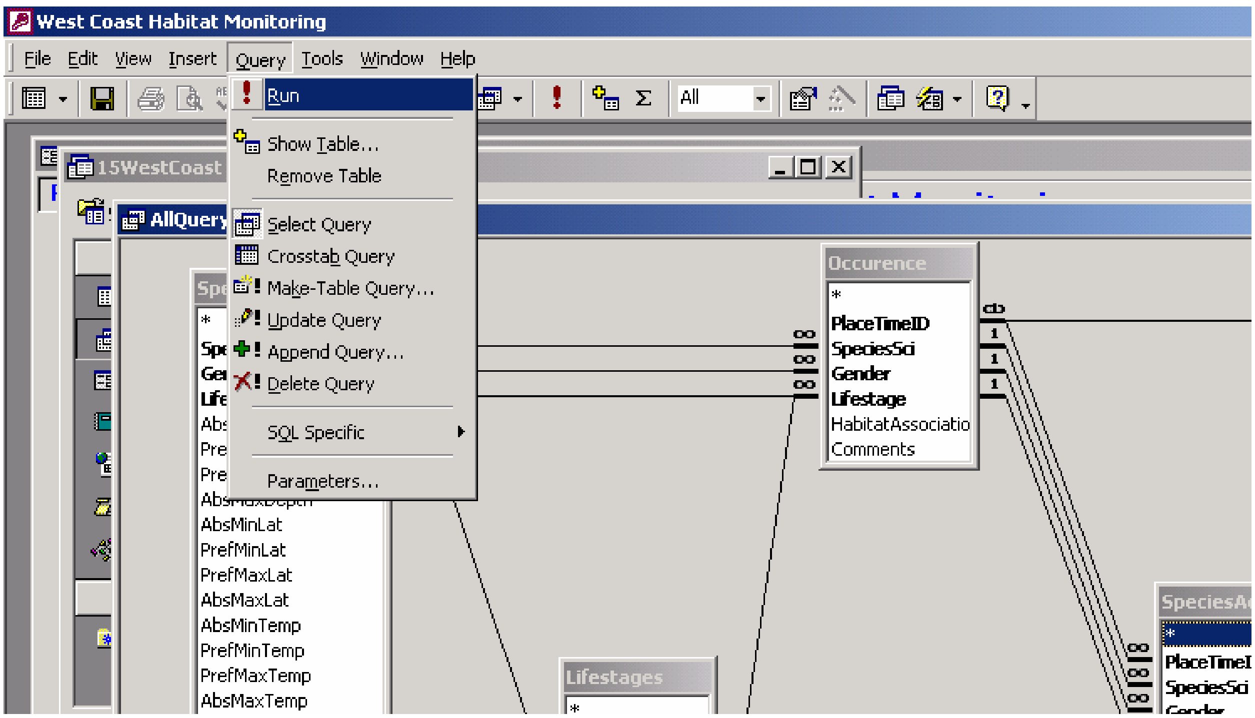


The 'foreign keys' are the lines representing the links between the tables. These are inherited from the relationships diagram. You can drag and size the tables to form the best layout.

You then double click or drag and drop the columns you require for your results. Note that a 'criterion' has been entered for the species name ('Coryphaenoides acrolepis') and that the SortOrder column of the Lifestages table has been utilized to make sure the results appear in life stage order,

You could copy this entire query to one of a new name and edit that to your preferences. You could add in or take out columns or conditions as you require. The simplest way to create your own new query is to open the AllQuery0 and use the 'File'/Save As…' option giving it the name of your choice.

You can run the query by a number of different methods the simplest being to press the red exclamation mark from the menu bar or toolbar.



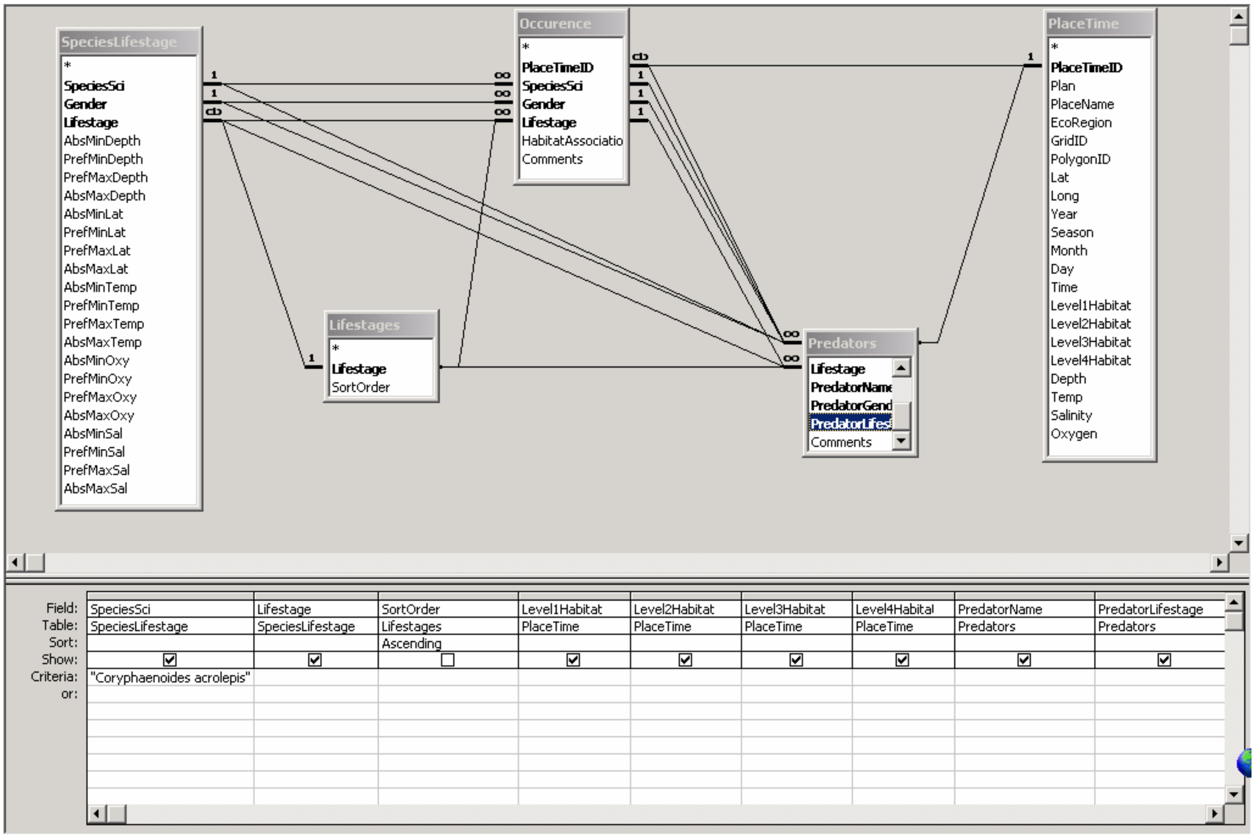
The results appear as follows:

AllQuery0 : Select Query

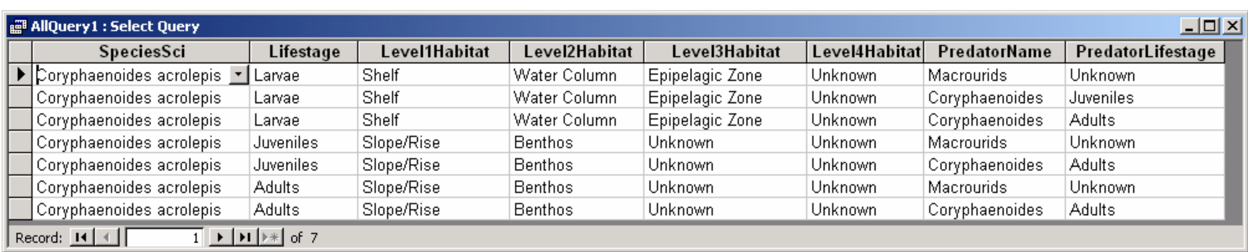
SpeciesSci	Lifestage	Level1Habitat	Level2Habitat	Level3Habitat	Level4Habitat	Activity	ActivityAssociation
Coryphaenoides acrolepis	Eggs	Shelf	Water Column	Epipelagic Zone	Unknown	Unknown	Unknown
Coryphaenoides acrolepis	Eggs	Slope/Rise/Plain	Water Column	Epipelagic Zone	Unknown	Unknown	Unknown
Coryphaenoides acrolepis	Larvae	Shelf	Water Column	Epipelagic Zone	Unknown	Feeding	Strong
Coryphaenoides acrolepis	Larvae	Slope/Rise/Plain	Water Column	Epipelagic Zone	Unknown	Feeding	Strong
Coryphaenoides acrolepis	Juveniles	Slope/Rise/Plain	Water Column	Mesopelagic Zone	Unknown	Feeding	Medium
Coryphaenoides acrolepis	Juveniles	Slope/Rise/Plain	Water Column	Epipelagic Zone	Unknown	Feeding	Medium
Coryphaenoides acrolepis	Juveniles	Slope/Rise	Benthos	Unknown	Unknown	Growth to M:	Strong
Coryphaenoides acrolepis	Juveniles	Slope/Rise	Benthos	Unknown	Unknown	Feeding	Strong
Coryphaenoides acrolepis	Adults	Slope/Rise	Benthos	Unconsolidated	Sand	All	Strong
Coryphaenoides acrolepis	Adults	Slope/Rise	Benthos	Unknown	Unknown	All	Strong
Coryphaenoides acrolepis	Adults	Slope/Rise	Basin	Unconsolidated	Sand	All	Strong

Record: 1 of 11

If instead you wished to investigate say predators then you would substitute the predators table for the activities table in the query design. See **'AllQuery1'**. That would appear as follows.



and the results would appear as follows:



SpeciesSci	Lifestage	Level1Habitat	Level2Habitat	Level3Habitat	Level4Habitat	PredatorName	PredatorLifestage
Coryphaenoides acrolepis	Larvae	Shelf	Water Column	Epipelagic Zone	Unknown	Macrourids	Unknown
Coryphaenoides acrolepis	Larvae	Shelf	Water Column	Epipelagic Zone	Unknown	Coryphaenoides	Juveniles
Coryphaenoides acrolepis	Larvae	Shelf	Water Column	Epipelagic Zone	Unknown	Coryphaenoides	Adults
Coryphaenoides acrolepis	Juveniles	Slope/Rise	Benthos	Unknown	Unknown	Macrourids	Unknown
Coryphaenoides acrolepis	Juveniles	Slope/Rise	Benthos	Unknown	Unknown	Coryphaenoides	Adults
Coryphaenoides acrolepis	Adults	Slope/Rise	Benthos	Unknown	Unknown	Macrourids	Unknown
Coryphaenoides acrolepis	Adults	Slope/Rise	Benthos	Unknown	Unknown	Coryphaenoides	Adults

The same substitution can be done for prey species and references.

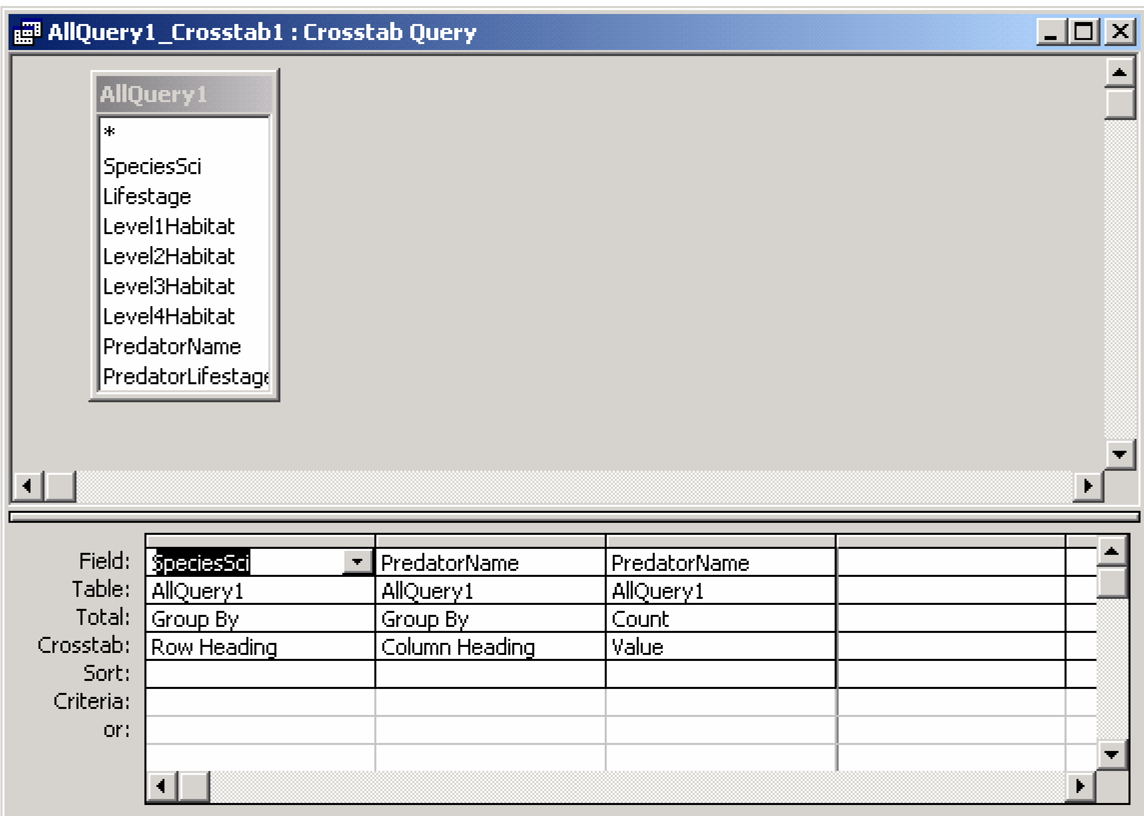
Such queries can form the source for other queries, charts that plot the results, or for exporting data to spreadsheets or other file formats for modeling etc.

One of the most common queries towards the end of a series of analyses is the cross-tab query which produces results more like a spreadsheet format. Indeed the results are often exported to spreadsheets for further manipulation.

For example, if we wished to find out which was the most common predator of a species regardless of the preys life stage or habitat setting we would form the following cross-tab query 'CrosstabAllQuery1' which takes the original query AllQuery1 as its input.

NB Remember that if you then alter such a source query you would then invalidate the cross-tab query based upon it. Care must be taken in this respect. It is often best to develop suites of parallel queries to avoid this pitfall and have some consistent naming conventions across and along the various streams to prevent the confusion that would otherwise develop.

From the main database window select the queries-new- Design View.
Select the Queries tab from the Show Table box and select AllQuery1.



Select the columns SpeciesSci once and PredatorName twice, should look as above.

Set the values in the grid as illustrated above.
Save the query as 'AllQuery1_Crosstab1' or whatever you wish to call it.

Run the query and the results will appear as follows.

revealing a total of three situations where Macrourids are the predators and 4 of cannibalism.

Note that if you remove the single species criterion from the source query AllQuery1 then you get the following results revealing what has had predator data entered and what those predators are (Column Headings), and what they eat.

SpeciesSci	Albacore	Anoplopoma f	Artedius harri	Atheresthes st	chaetognaths	Clupeids	Coryphaenoides acrolepis	Crabs
Anoplopoma fimbria								
Coryphaenoides acrolepis							4	
Eopsetta jordani								
Gadus macrocephalus								
Galeorhinus zyopterus								
Hexagrammos decagrammu								
Hydrolagus colliei								
Lepidopsetta bilineata								
Merluccius productus	2	2				3		
Microstomus pacificus	3	4						
Ophiodon elongatus								
Platichthys stellatus								
Pleuronectes vetulus								
Psettichthys melanostictus								
Scorpaenichthys marmoratu			1					
Sebastes alutus		2		2				
Sebastes atrovirens					1			
Sebastes auriculatus								
Sebastes chrysomelas					1			
Sebastes crameri	4							
Sebastes dalli					1			
Sebastes jordani								
Sebastes maliger					2			
Sebastes melanops					2			
Sebastes nebulosus					1			
Sebastes nigrocinctus					2			
Sebastes paucispinis	2							
Sebastes rastrelliger								
Sebastes serranoides					2			
Sebastes serriceps								
Sebastolobus alascanus								
Sebastolobus altivelis		1						

You could specify multiple criteria for both species and predator species to limit the results set depending on your line of investigation. The same kind of investigations could be made for Activities or prey data and all could be further refined by select only some levels of habitat classifications and only certain values within these. You could look at say only level2 habitats and only ‘benthos’ from within these.

You can further refine queries by editing the ‘SQL’ code version of them. This is particularly useful when creating more elaborate cross-tab queries and filters for charts etc.

You can select the SQL view from under View on the menu bar. The SQL for the AllQuery1 query would look like the following:

SELECT SpeciesLifestage.SpeciesSci, SpeciesLifestage.Lifestage, PlaceTime.Level1Habitat, PlaceTime.Level2Habitat, PlaceTime.Level3Habitat, PlaceTime.Level4Habitat, Predators.PredatorName, Predators.PredatorLifestage

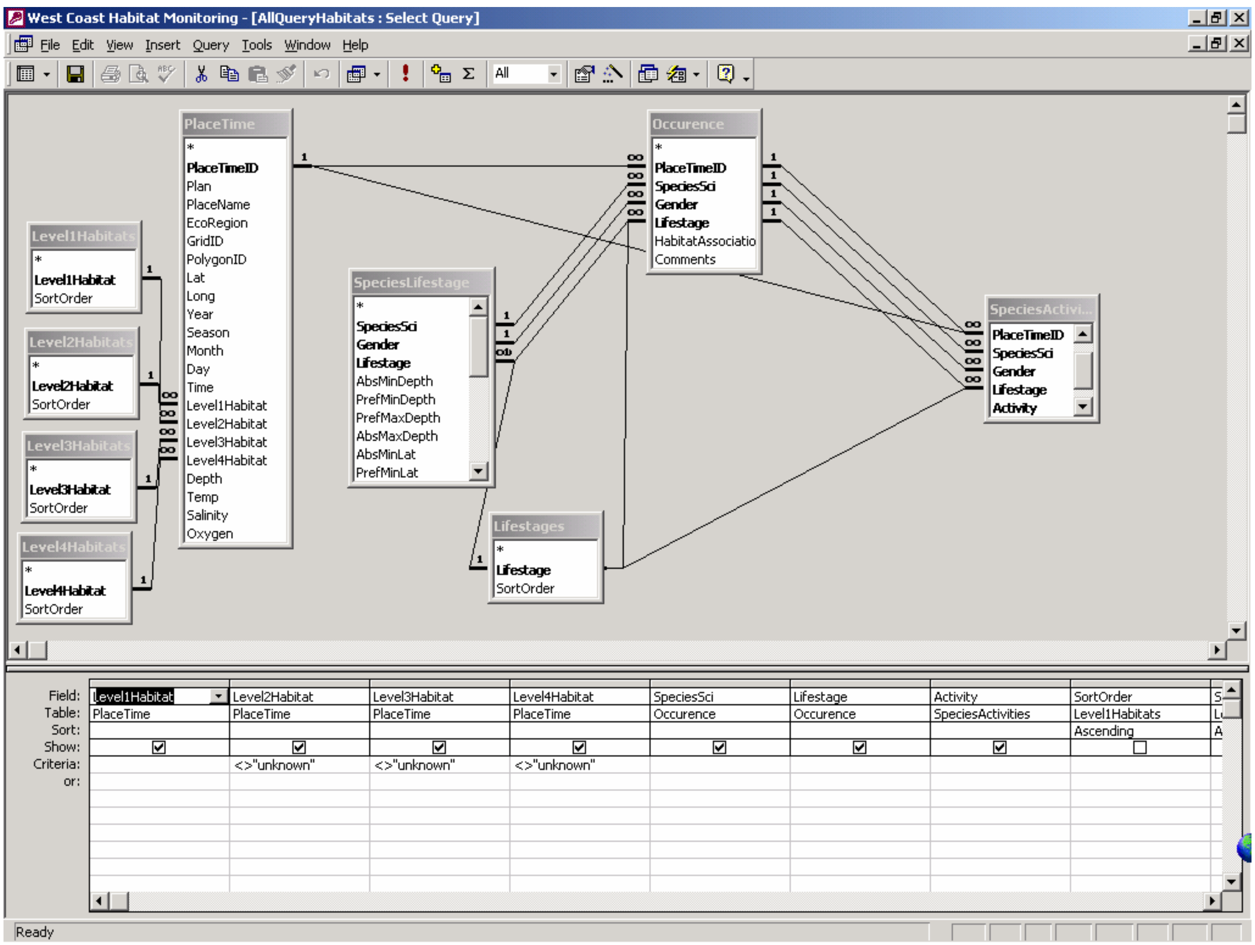
```
FROM ((Occurence INNER JOIN Lifestages ON Occurence.Lifestage = Lifestages.Lifestage)
INNER JOIN (PlaceTime INNER JOIN Predators ON PlaceTime.PlaceTimeID =
Predators.PlaceTimeID) ON (PlaceTime.PlaceTimeID = Occurence.PlaceTimeID) AND
(Occurence.Lifestage = Predators.Lifestage) AND (Occurence.Gender = Predators.Gender) AND
(Occurence.SpeciesSci = Predators.SpeciesSci) AND (Occurence.PlaceTimeID =
Predators.PlaceTimeID) AND (Lifestages.Lifestage = Predators.Lifestage)) INNER JOIN
SpeciesLifestage ON (SpeciesLifestage.Lifestage = Predators.Lifestage) AND
(SpeciesLifestage.Gender = Predators.Gender) AND (SpeciesLifestage.SpeciesSci =
Predators.SpeciesSci) AND (SpeciesLifestage.Lifestage = Occurence.Lifestage) AND
(SpeciesLifestage.Gender = Occurence.Gender) AND (SpeciesLifestage.SpeciesSci =
Occurence.SpeciesSci) AND (Lifestages.Lifestage = SpeciesLifestage.Lifestage)
WHERE (((SpeciesLifestage.SpeciesSci)="Coryphaenoides acrolepis"))
ORDER BY Lifestages.SortOrder;
```

3.4.1.3 Example 2: Habitat based investigation

This shorter example, 'AllQueryHabitats', demonstrates how to develop similar lines of queries except they are based on the perspective of habitats rather than species.

It is probably best to first look at the data from the 'Place-Time centric' data form (exactly the same data but arranged from a habitats perspective), which is chosen from the main opening form or from the drop down menu which is part of the West Coast Tools tool bar.

This query orders all habitats according to the values within the four habitat levels and assumes each level is nested within the previous. Then for each unique combination of habitats it lists the species life stages and their activities.



A portion of the results appear as follows:

Level1Habitat	Level2Habitat	Level3Habitat	Level4Habitat	SpeciesSci	Lifestage	Activity
Estuarine	Intertidal Benthos	Unconsolidated	Mud	Squalus acanthias	Juveniles	Growth to Maturity
Estuarine	Intertidal Benthos	Unconsolidated	Mud	Squalus acanthias	Adults	All
Estuarine	Intertidal Benthos	Unconsolidated	Mud	Raja inornata	Adults	All
Estuarine	Intertidal Benthos	Unconsolidated	Mud	Triakis semifasciata	Adults	All
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Hippoglossoides elassodon	Juveniles	Feeding
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Errex zachirus	Juveniles	Feeding
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Eopsetta jordani	Juveniles	Feeding
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Platichthys stellatus	Juveniles	Feeding
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Psettichthys melanostictus	Juveniles	Growth to Maturity
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Lepidopsetta bilineata	Juveniles	Growth to Maturity
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Psettichthys melanostictus	Adults	All
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Platichthys stellatus	Adults	All
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Lepidopsetta bilineata	Adults	All
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Gadus macrocephalus	Adults	Feeding
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Errex zachirus	Adults	Feeding
Estuarine	Subtidal Benthos	Unconsolidated	Mixed mud/sand	Hippoglossoides elassodon	Adults	All
Estuarine	Subtidal Benthos	Unconsolidated	Sand	Lepidopsetta bilineata	Eggs	Unknown
Estuarine	Subtidal Benthos	Unconsolidated	Mud	Raja inornata	Eggs	Unknown
Estuarine	Subtidal Benthos	Unconsolidated	Sand	Microstomus pacificus	Juveniles	Growth to Maturity
Estuarine	Subtidal Benthos	Unconsolidated	Sand	Ophiodon elongatus	Juveniles	Feeding

Note that results have been filtered out where habitat values are ‘Unknown,’ which in fact is the majority of cases.

Again as with the first example, this query can be copied and renamed and then used as a template to extend and vary it, develop cross tabs and charts etc.

As with the first example from the species perspective, the SpeciesActivities table could be substituted with the Predators, Prey or references tables and the query modified to analyze the attributes in these tables instead.

3.4.1.4 Example 3: Using species level attributes

This example demonstrates an analysis of the general attributes recorded at the species level, i.e. absolute and preferred ranges of latitude, depth, temperature, salinity or oxygen.

You will have noted from the data entry screens that these attributes can be recorded at two levels of detail

a) the general ranges associated with a species, and

b) more precise values associated with a particular ‘TimePlaceID.’

How precise would depend on what level of detail is used with the PlaceTimeID. It could be a period for an area or a specific location at one exact time or for a given habitat definition. At present (October 2003) none of the facilities offered for the more detailed recording offered with b) are required or being utilized. Thus only the general species wide values are being used and it follows that these have to be applied to all locations and habitat types for all times. This affects the way the query is structured with those physiographic attributes being sourced from the SpeciesLifestage table.

In fact there is very little variation in the values of the extracted data for depth and latitude for each species; hence the results tend to be somewhat ‘uninteresting’ in terms of the variety in the

query and chart output used to demonstrate the potential of these kinds of queries. Thus for the purposes of the best demonstration of the potential of these kinds of queries, the species, *Merluccius productus*, with most detailed variety has been chosen and the somewhat crude mid points of absolute ranges of latitude and depth used.

The final charted output appears as follows (Figure 13):

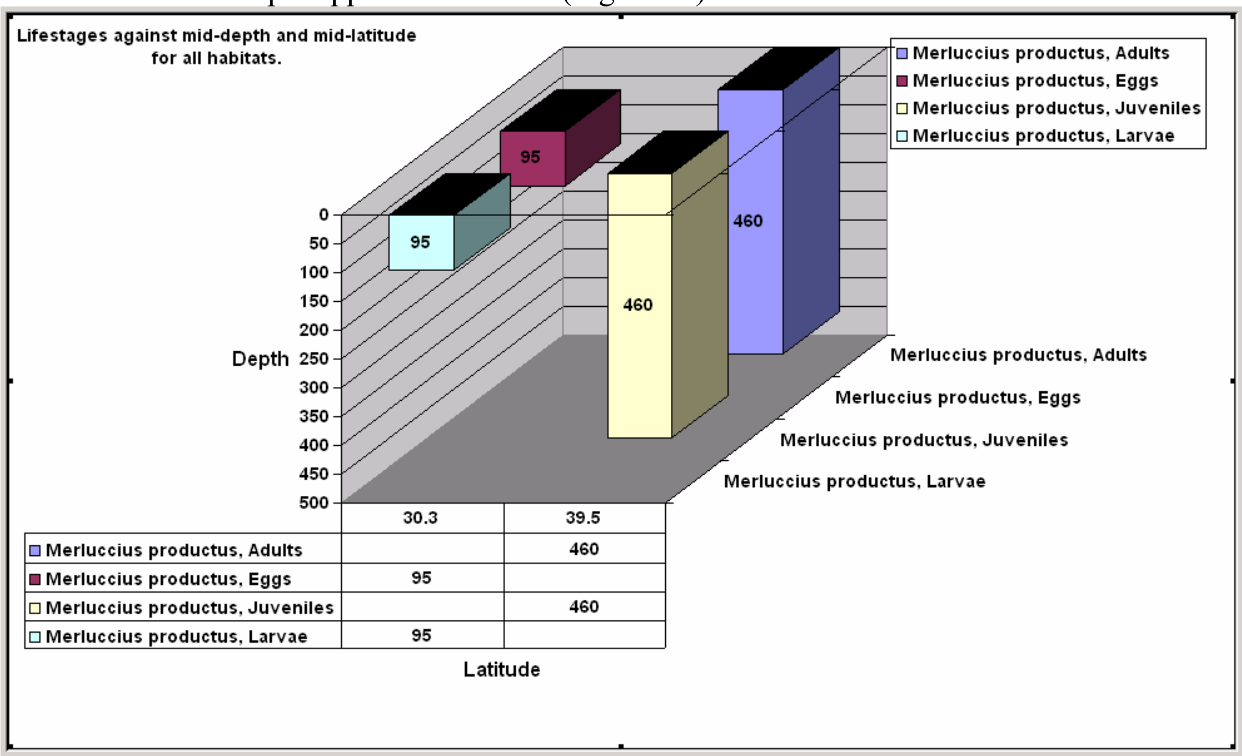


Figure 4 Life stages against mid depth and mid latitude from chart 'chtLifestageLatDepth'

This is one of the few species with where there is enough variety in the extracted depth/ latitude data to demonstrate the range of possible plotting. Most other species have the same depths and/or latitudes for each of the life stages.

This chart is also available from the main opening form under the Charts section via the button 'Lifestages by Lat-Depth'

The results of the underlying select query (which is named 'LifestageLatDepth') looks like this:

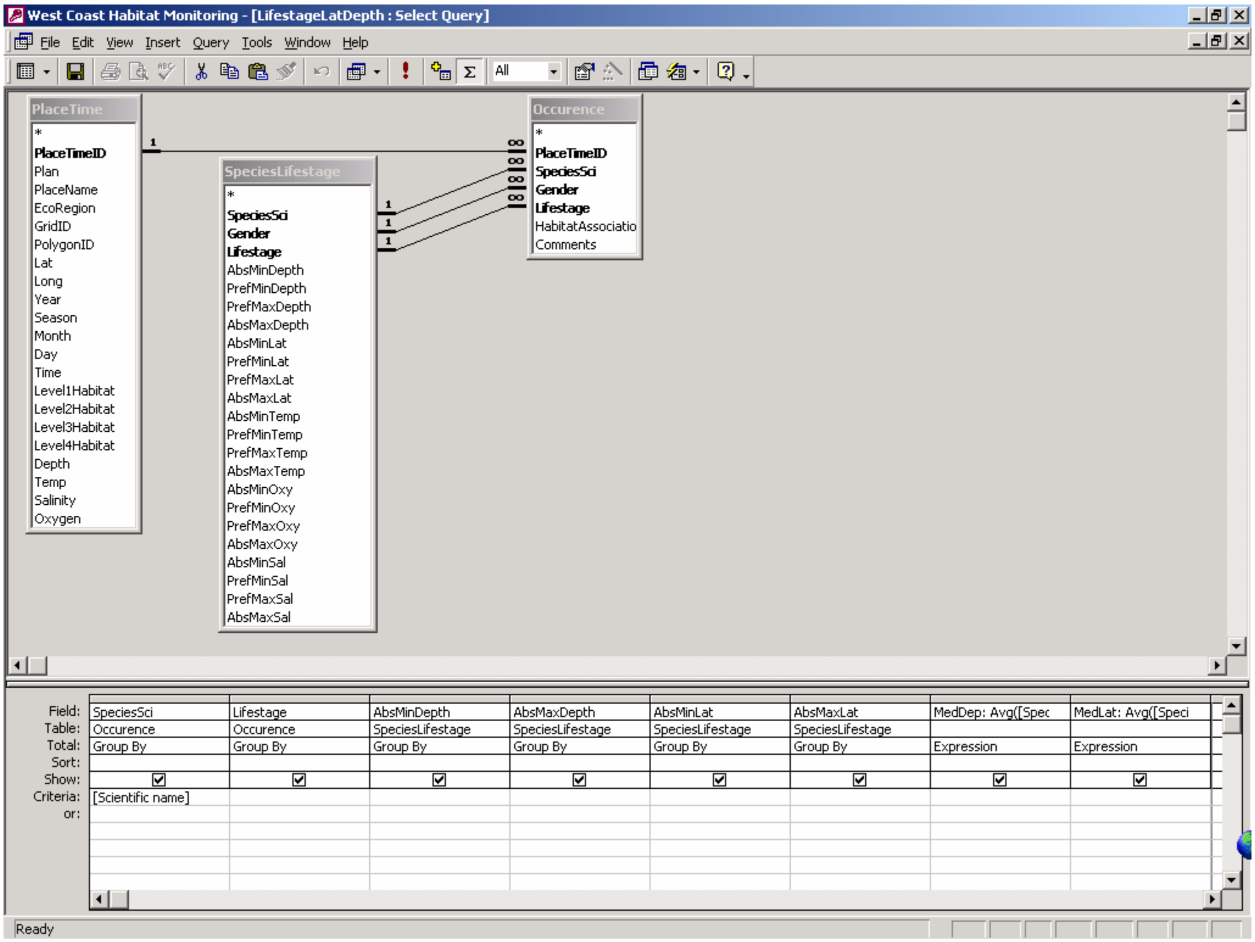
LifestageLatDepth : Select Query

SpeciesSci	Lifestage	AbsMinDepth	AbsMaxDepth	AbsMinLat	AbsMaxLat	MedDep	MedLat
Merluccius productus	Adults	0	920	24.5	54.5	460	39.5
Merluccius productus	Eggs	40	150	24.5	36	95	30.25
Merluccius productus	Juvenile	0	920	24.5	54.5	460	39.5
Merluccius productus	Larvae	40	150	24.5	36	95	30.25

Record: 1 of 4

This query is also available via the button ‘Lifestages by Lat-Depth’ under the Queries section on the main opening form ‘frmMain’.

The query is structures as follows:



This also demonstrates the use of ‘expressions’ in queries. The ‘MedDep’ and ‘MedLat’ in the query. This is where an output field is based on an underlying calculation rather than a simple value or simple aggregate function of those values (a straight average, count or sum etc)

Again you can copy and rename this query and use that new copy as a template to alter and develop your own queries.

3.4.1.5 Example 4: Counts of Habitats

There are a series of other queries and charts that provide examples of how aggregate functions can be used in Access. Again the form of the examples used is more to demonstrate what is possible within Access rather than for biological analytical rigor! User can use these examples as templates to develop their own queries that are appropriate to their line of biological investigation.

The query 'Habitats by Species Lifestage' simply counts the occurrence of sub-habitats within each of the level1 habitats, for each of the SpeciesLifestages.

Habitats by Species Lifestage : Select Query

SpeciesSci	Lifestage	Level1Habitat	CountOfLevel1Habitat
Anoplopoma fimbria	Adults	Shelf	2
Anoplopoma fimbria	Adults	Slope/Rise	1
Anoplopoma fimbria	Eggs	Slope/Rise	2
Anoplopoma fimbria	Juveniles	Shelf	1
Anoplopoma fimbria	Juveniles	Slope/Rise	3
Anoplopoma fimbria	Larvae	Shelf	1
Anoplopoma fimbria	Larvae	Slope/Rise	6
Antimora microlepis	Adults	Shelf	1
Antimora microlepis	Adults	Slope/Rise	1
Atheresthes stomias	Adults	Shelf	7

The query is also available from the main opening form under the queries section via the button 'Habitats by Species Lifestage'.

You can look at it in design view to see its simple structure.

This query is used as source data for the following query 'XtabBySpeciesByHabitat' which simply cross-tabs the output using the level1 habitat values as column headings instead of leaving them as row headings. That query is also available via the button 'Crosstab species by Habitat1' on the main opening form.

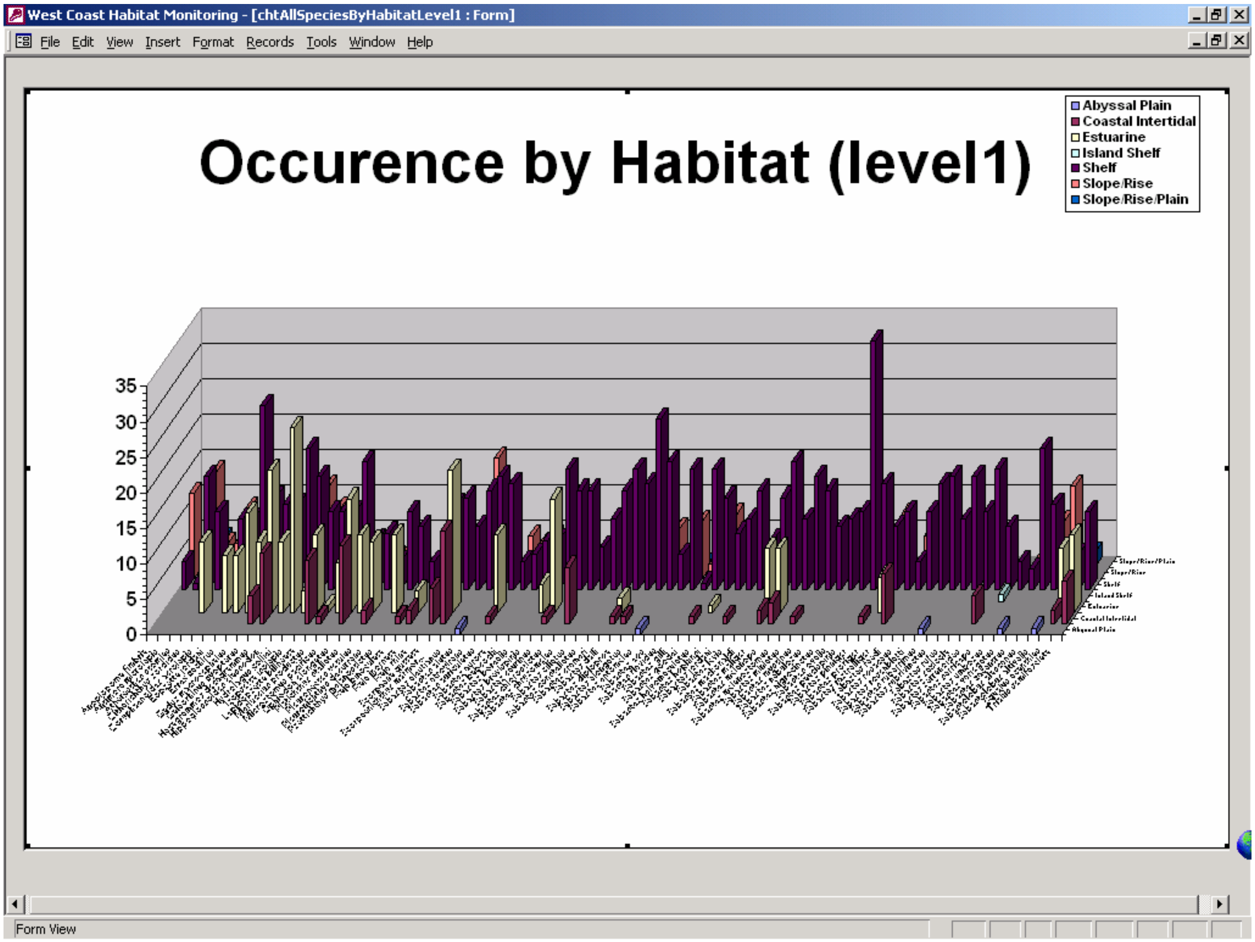
XtabSpeciesByHabitat : Crosstab Query

SpeciesSci	Abyssal Plain	Coastal Interti	Estuarine	Island Shelf	Shelf	Slope/Rise	Slope/Rise/Plain
Anoplopoma fimbria					4	12	
Antimora microlepis					1	1	
Atheresthes stomias					16	15	4
Citharichthys sordidu			10		11	5	2
Coryphaenoides acrc					2	5	4
Eopsetta jordani			8		10	10	1
Errex zachirus			8		10	8	2
Gadus macrocephalu			14		26	6	1
Galeorhinus zyopteru		4	10		13		
Hexagrammos decag		10	20		12		2
Hippoglossoides elas			10		12	8	
Hydrolagus colliei			26		20		
Isopsetta isolepis			3		16	13	

Record: 1 of 81

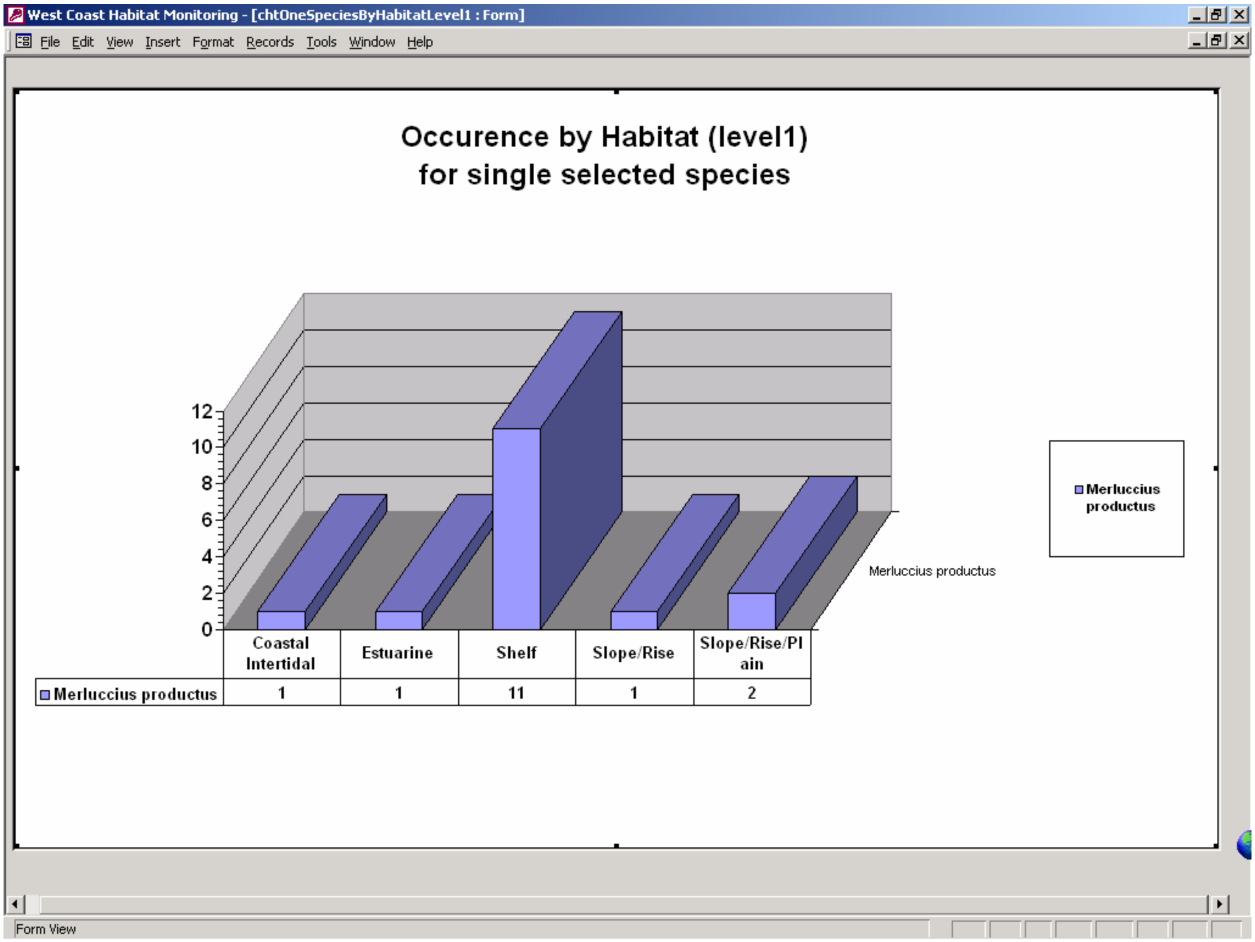
The same principle of cross-tabbing habitats by species life stage is used as the source for the two charts.

The chart ‘chtAllSpeciesByHabitatLevel1’ which counts and charts level 1 sub-habitat for all the species looks as follows:



This chart is also available from the main form via the button ‘AllSpeciesByHab.Level1’.

There is also a chart ‘chtOneSpeciesByHabitatLevel1’ which does the same but, for a single species for which the user supplies the name as a parameter.



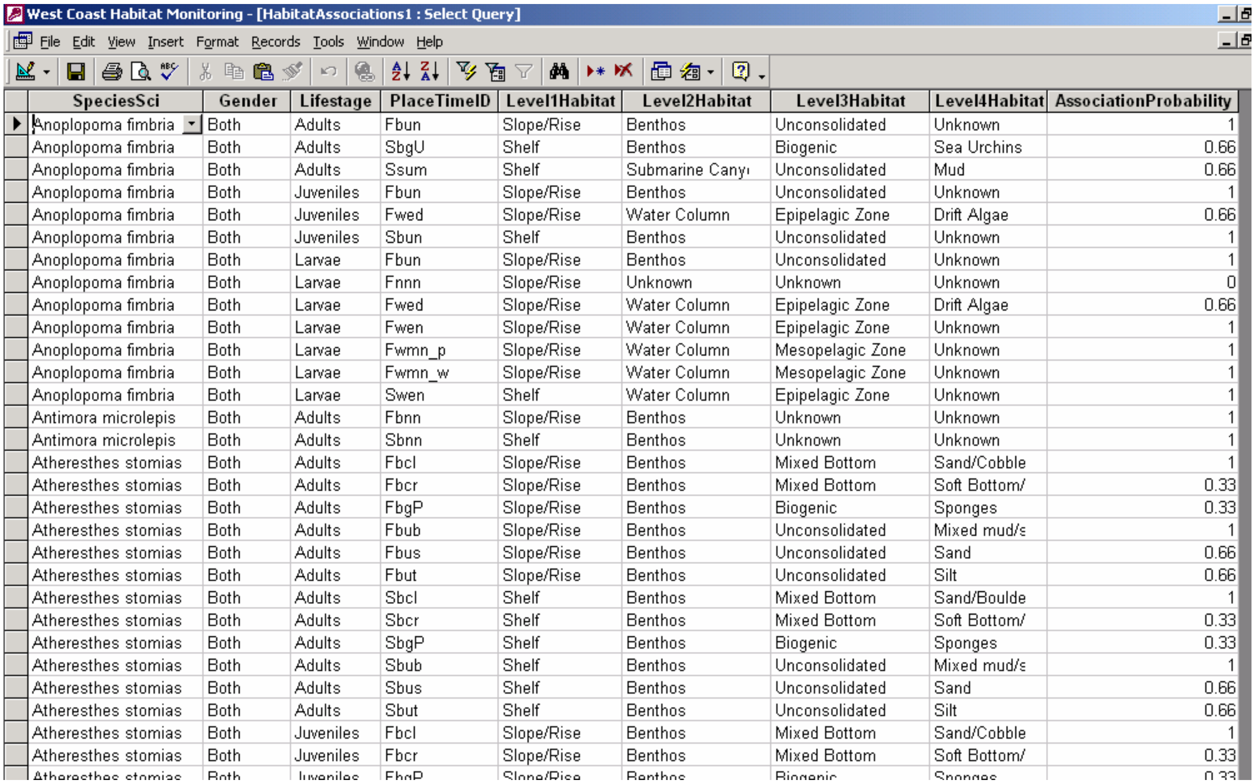
This chart is also available from the main form via the button 'OneSpeciesByHab.Level1'.

These charts are constructed within forms. Their cross-tabbed data sources are specified as SQL clauses under their 'properties.'

3.4.1.6 Example 5: Data extraction

The query 'HabitatAssociations1' was used to extract an assemblage of data required for the Bayesian modeling software.

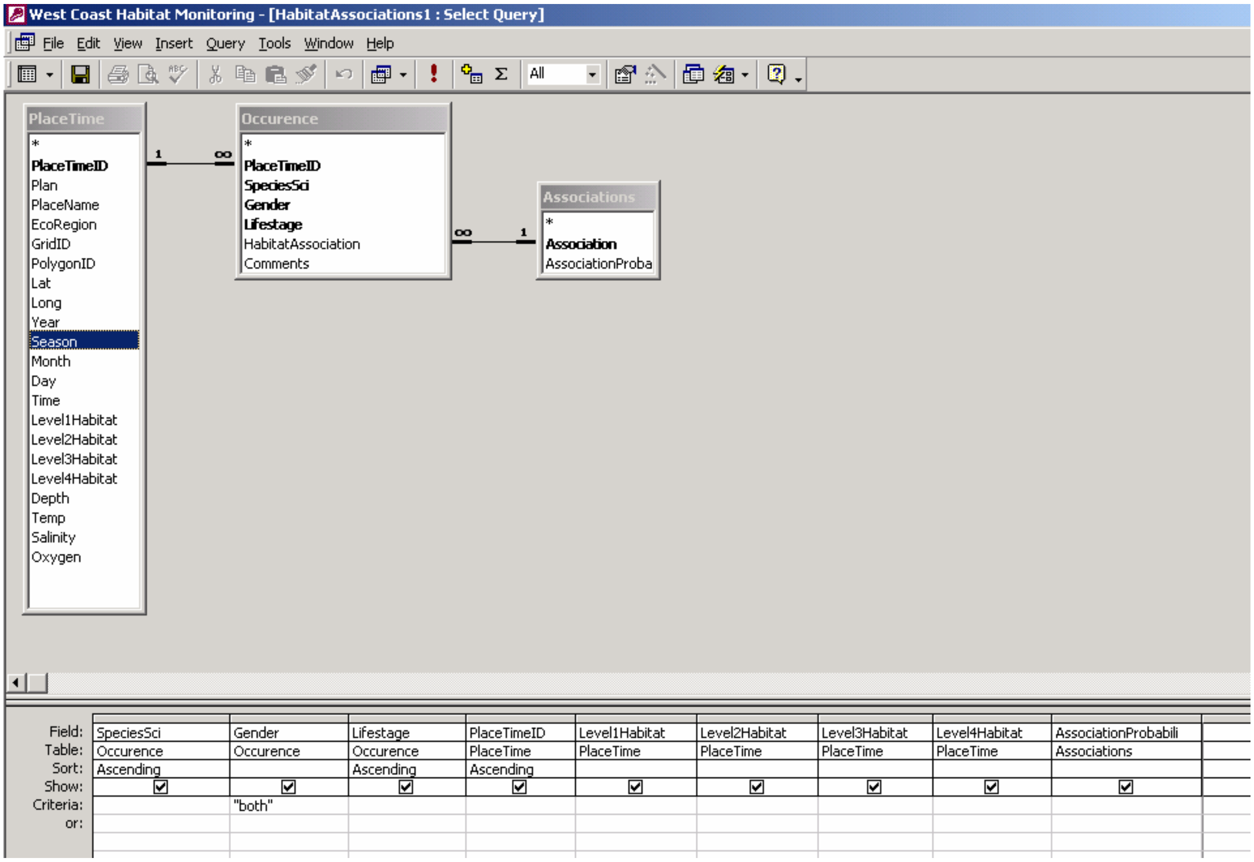
A portion of the output appears as follows:



SpeciesSci	Gender	Lifestage	PlaceTimeID	Level1Habitat	Level2Habitat	Level3Habitat	Level4Habitat	AssociationProbability
Anoplopoma fimbria	Both	Adults	Fbun	Slope/Rise	Benthos	Unconsolidated	Unknown	1
Anoplopoma fimbria	Both	Adults	SbgU	Shelf	Benthos	Biogenic	Sea Urchins	0.66
Anoplopoma fimbria	Both	Adults	Ssum	Shelf	Submarine Cany	Unconsolidated	Mud	0.66
Anoplopoma fimbria	Both	Juveniles	Fbun	Slope/Rise	Benthos	Unconsolidated	Unknown	1
Anoplopoma fimbria	Both	Juveniles	Fwed	Slope/Rise	Water Column	Epipelagic Zone	Drift Algae	0.66
Anoplopoma fimbria	Both	Juveniles	Sbun	Shelf	Benthos	Unconsolidated	Unknown	1
Anoplopoma fimbria	Both	Larvae	Fbun	Slope/Rise	Benthos	Unconsolidated	Unknown	1
Anoplopoma fimbria	Both	Larvae	Fnnn	Slope/Rise	Unknown	Unknown	Unknown	0
Anoplopoma fimbria	Both	Larvae	Fwed	Slope/Rise	Water Column	Epipelagic Zone	Drift Algae	0.66
Anoplopoma fimbria	Both	Larvae	Fwen	Slope/Rise	Water Column	Epipelagic Zone	Unknown	1
Anoplopoma fimbria	Both	Larvae	Fwmn_p	Slope/Rise	Water Column	Mesopelagic Zone	Unknown	1
Anoplopoma fimbria	Both	Larvae	Fwmn_w	Slope/Rise	Water Column	Mesopelagic Zone	Unknown	1
Anoplopoma fimbria	Both	Larvae	Swen	Shelf	Water Column	Epipelagic Zone	Unknown	1
Antimora microlepis	Both	Adults	Fbnn	Slope/Rise	Benthos	Unknown	Unknown	1
Antimora microlepis	Both	Adults	Sbnn	Shelf	Benthos	Unknown	Unknown	1
Atheresthes stomias	Both	Adults	Fbcl	Slope/Rise	Benthos	Mixed Bottom	Sand/Cobble	1
Atheresthes stomias	Both	Adults	Fbcr	Slope/Rise	Benthos	Mixed Bottom	Soft Bottom/	0.33
Atheresthes stomias	Both	Adults	FbgP	Slope/Rise	Benthos	Biogenic	Sponges	0.33
Atheresthes stomias	Both	Adults	Fbub	Slope/Rise	Benthos	Unconsolidated	Mixed mud/s	1
Atheresthes stomias	Both	Adults	Fbus	Slope/Rise	Benthos	Unconsolidated	Sand	0.66
Atheresthes stomias	Both	Adults	Fbut	Slope/Rise	Benthos	Unconsolidated	Silt	0.66
Atheresthes stomias	Both	Adults	Sbcl	Shelf	Benthos	Mixed Bottom	Sand/Boulde	1
Atheresthes stomias	Both	Adults	Sbcr	Shelf	Benthos	Mixed Bottom	Soft Bottom/	0.33
Atheresthes stomias	Both	Adults	SbgP	Shelf	Benthos	Biogenic	Sponges	0.33
Atheresthes stomias	Both	Adults	Sbub	Shelf	Benthos	Unconsolidated	Mixed mud/s	1
Atheresthes stomias	Both	Adults	Sbus	Shelf	Benthos	Unconsolidated	Sand	0.66
Atheresthes stomias	Both	Adults	Sbut	Shelf	Benthos	Unconsolidated	Silt	0.66
Atheresthes stomias	Both	Juveniles	Fbcl	Slope/Rise	Benthos	Mixed Bottom	Sand/Cobble	1
Atheresthes stomias	Both	Juveniles	Fbcr	Slope/Rise	Benthos	Mixed Bottom	Soft Bottom/	0.33
Atheresthes stomias	Both	Juveniles	FbgP	Slope/Rise	Benthos	Biogenic	Sponges	0.33

It basically lists the numeric probability of habitat association for all habitats for each SpeciesLifestage and gender.

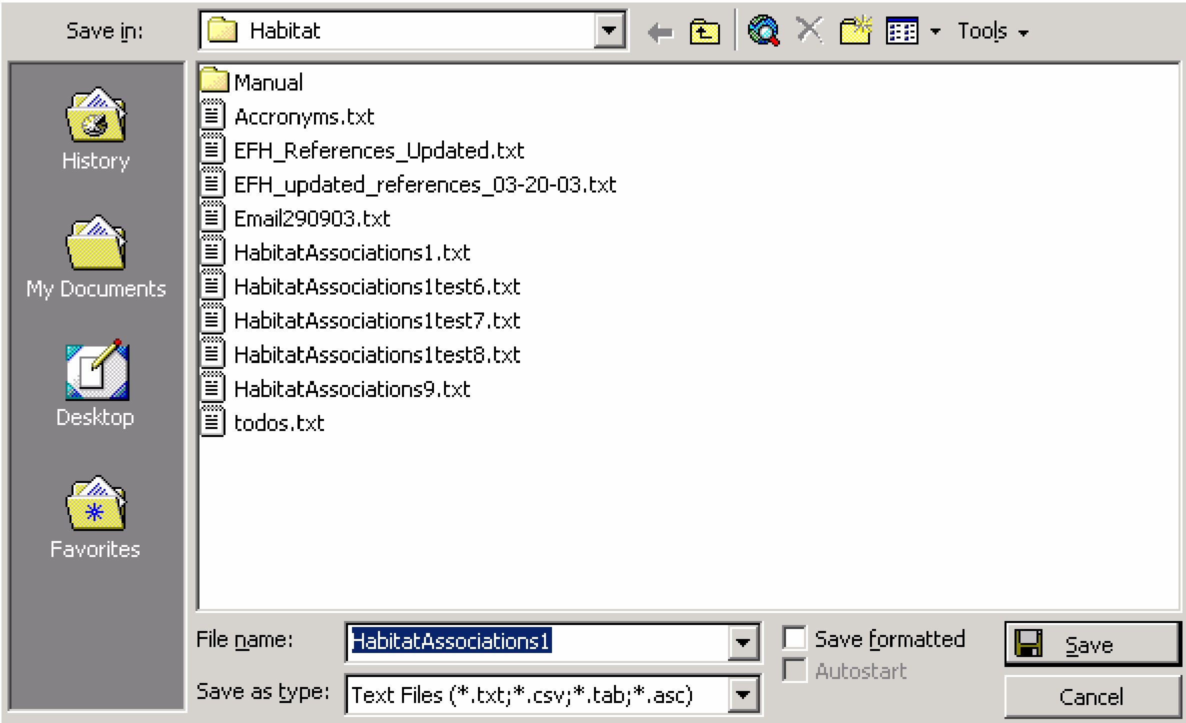
Its design view looks like this:



Once the query is written, it is used as a source for a menu-driven export routine that can export the data in a wide range of formats including that of spreadsheet files and standard comma delimited text files. The choice of format depends on the receiving software.

You can export a datasheet to a delimited or fixed-width text file; to do this, in the Database window, click the name of the table, query, view, or stored procedure you want to export, and then on the File menu, click Export.

The following screen comes up:



In the “Save as type” box, click Text Files (*.txt;*.csv;*.tab;*.asc).

Click the arrow to the right of the Save In box, and select the drive or folder to export to.

In the File Name box, enter the file name, and then click Save.

NB Make sure the ‘Save formatted’ box is NOT ticked.

Microsoft Access then starts the Export Text Wizard.

Follow the directions in the dialog boxes. Click Advanced to create or use an import/export specification.

You can call up this specification for re-use in future should you repeat the export procedure. You still have to go through the menu system but it at least remembers the settings you previously specified. It is also possible to save an export specification as a macro or visual basic code module. This can be done if required though for the assumed usage here we confine ourselves to the menu system which is powerful, flexible, and easy to use.

The resultant text appears as follows:

SpeciesSci,Gender,Lifestage,PlaceTimeID,Level1Habitat,Level2Habitat,Level3Habitat,Level4Habitat,AssociationProbability
Anoplopoma fimbria,Both,Adults,Fbun,Slope/Rise,Benthos,Unconsolidated,Unknown,100

Anoplopoma fimbria,Both,Adults,SbgU,Shelf,Benthos,Biogenic,Sea Urchins,66
Anoplopoma fimbria,Both,Adults,Ssum,Shelf,Submarine Canyon,Unconsolidated,Mud,66
Anoplopoma fimbria,Both,Juveniles,Fbun,Slope/Rise,Benthos,Unconsolidated,Unknown,100
Anoplopoma fimbria,Both,Juveniles,Fwed,Slope/Rise,Water Column,Epipelagic Zone,Drift Algae,66
Anoplopoma fimbria,Both,Juveniles,Sbun,Shelf,Benthos,Unconsolidated,Unknown,100
Anoplopoma fimbria,Both,Larvae,Fbun,Slope/Rise,Benthos,Unconsolidated,Unknown,100
Anoplopoma fimbria,Both,Larvae,Fnnn,Slope/Rise,Unknown,Unknown,Unknown,0
Anoplopoma fimbria,Both,Larvae,Fwed,Slope/Rise,Water Column,Epipelagic Zone,Drift Algae,66
Anoplopoma fimbria,Both,Larvae,Fwen,Slope/Rise,Water Column,Epipelagic Zone,Unknown,100
Anoplopoma fimbria,Both,Larvae,Fwmn_p,Slope/Rise,Water Column,Mesopelagic Zone,Unknown,100
Anoplopoma fimbria,Both,Larvae,Fwmn_w,Slope/Rise,Water Column,Mesopelagic Zone,Unknown,100
Anoplopoma fimbria,Both,Larvae,Swen,Shelf,Water Column,Epipelagic Zone,Unknown,100
Antimora microlepis,Both,Adults,Fbnn,Slope/Rise,Benthos,Unknown,Unknown,100
Antimora microlepis,Both,Adults,Sbnn,Shelf,Benthos,Unknown,Unknown,100
Atheresthes stomias,Both,Adults,Fbcl,Slope/Rise,Benthos,Mixed Bottom,Sand/Cobble,100
Atheresthes stomias,Both,Adults,Fbcr,Slope/Rise,Benthos,Mixed Bottom,Soft Bottom/rock,33

Etc etc …

APPENDIX 11A. EXAMPLE DATA EXTRACTION FROM UPDATED LIFE HISTORY DESCRIPTIONS

CANARY ROCKFISH (*Sebastes pinniger*)

Range

Canary rockfish are found between Cape Colnett, Baja California, and southeastern Alaska (lat. 56°N, long. 134°W) (Boehlert 1980, Boehlert and Kappenman 1980, Hart 1973, Love 1996, Miller and Lea 1972, Richardson and Laroche 1979).

Fishery

Canary rockfish are a major constituent of the commercial trawl fishery off Oregon and Washington (Boehlert 1980, Gunderson and Lenarz 1980, Love 1996). Off California, canary rockfish are caught mainly in the sport and commercial longline fisheries. They are moderately important in the party and private vessel sport fishery, from central California northward (Boehlert 1980, Love 1996).

Habitat

Canary rockfish are considered a middle shelf-mesobenthal species (Allen and Smith 1988). There is a major population concentration of canary rockfish between latitude 44° 30' and 45° 00' N off Oregon (Richardson and Laroche 1979).

Canary rockfish have a depth range from the surface (juveniles) to 274 m (Boehlert 1980, Hart 1973, Love 1996), but primarily inhabit waters 91-183 m deep (Boehlert and Kappenman 1980). Larvae and juveniles are pelagic (Boehlert and Kappenman 1980, Richardson and Laroche 1979). Larvae can be captured over a wide area, from 13-306 km offshore, and pelagic juveniles occur mostly beyond the continental shelf (Richardson and Laroche 1979).

Canary rockfish inhabit shallow water when they are young and deep water as adults (Mason 1995). Adults have two primary habitat preferences: some are semipelagic, forming loose schools above rocky areas; and some are nonschooling, solitary benthic individuals (Stein et al. 1992). Adult canary rockfish are associated with pinnacles and sharp drop-offs (Love 1996). They are also found near, but usually not on the bottom, often associating with yellowtail, widow, and silvergray rockfish (Love 1996). Canary rockfish are most abundant above hard bottoms (Boehlert and Kappenman 1980), and they have been observed among mixtures of mud and boulders (Love et al. 2002). In the southern part of its range, the canary rockfish appears to be a reef-associated species (Boehlert 1980). On Heceta Bank, near Oregon, they were commonly found in boulder and cobble fields in association with rosethorn, sharpchin, yelloweye and pygmy rockfish (Stein et al. 1992). In studies conducted off Southeast Alaska using an ROV, Johnson et al. (2003) reported finding canary rockfish primarily associated with

complex bottoms composed of rocks and boulders, and a few individuals were seen near soft sediments.

Young-of-the-year rockfish can also be found in tide pools (Love 1996), and are associated with artificial reefs, and in interfaces between mud and rock (Cailliet et al. 2000). In central California, young-of-the-year (YOY) canary rockfish are first observed near the bottom at the seaward, sand-rock interface and farther seaward in deeper water (18-24 m) (Carr 1991). Their first appearance generally occurs shortly after the first upwellings of the spring (Carr 1991). They are often seen hovering above sand or small rock piles (VenTresca et al. 1996), and are seldom associated with kelp beds, although some YOY are associated with floating algae (Carr 1991).

Migrations and Movements

Canary rockfish are densely aggregating fish (Love 1996). Juveniles descend into deeper water as they mature (Love 1996). Canary rockfish move into deeper water with age and also are capable of major latitudinal movements (up to 380 nautical miles) (Lea et al. 1999). Juveniles have been reported to be associated with rocky sandy areas during the day and with sand flats during the night (Love et al. 2002).

Reproduction

Canary rockfish are ovoviviparous and have internal fertilization (Boehlert and Kappenman 1980, Richardson and Laroche 1979). Off California, canary rockfish spawn from November-March and from January-March off Oregon, Washington, and British Columbia (Hart 1973, Love 1996, Richardson and Laroche 1979). A wide range in larval sizes over a broad time span indicates that canary rockfish may have protracted and variable spawning (Richardson and Laroche 1979).

The age of 50% maturity of canary rockfish is 9 years; nearly all are mature by age 13 (Paul Reilly, personal communication). Maximum age has been estimated as 60 years (Adams 1992) to 75 years (ODFW, personal communication).

Growth and Development

The mean length of newly extruded canary rockfish larvae is 3.66 mm SL (Richardson and Laroche 1979). The transformation to pelagic juvenile occurs at sizes greater than 12.5 mm SL. Transformation to benthic juveniles occurs after 59.4 mm, during June-August (Richardson and Laroche 1979). Canary rockfish growth does not vary with latitude (Boehlert and Kappenman 1980). The maximum length canary rockfish grow to is 76 cm (Boehlert and Kappenman 1980, Hart 1973, Love 1996).

Off California, about 50% of the population is mature at 35.6 cm (5 or 6 years). A 48.3-cm long female carries approximately 260,000 young and fish 53.3- to 66-cm long carries about 1,900,000 young (Hart 1973). Canary rockfish can live to be 75 years old. A 10-year-old canary rockfish is approximately 50 cm SL (Love 1996). After age 11, females grow faster than males

and mature at a larger size, but males live longer (Boehlert 1980, Boehlert and Yoklavich 1984, Love 1996).

Trophic Interactions

Canary rockfish primarily prey on planktonic creatures, such as krill, and occasionally on fish (Love 1996). Canary rockfish feeding increases during the spring-summer upwelling period when euphausiids are the dominant prey and the frequency of empty stomachs is lower (Boehlert et al. 1989).

APPENDIX 11B

List of tables:

Name
Activities
Associations
EcoRegions
Grids
Influences
Level1Habitats
Level2Habitats
Level3Habitats
Lifestages
Occurence
OtherActivities
phablist
PlaceTime
Plans
Predators
Prey
ReferenceInstance
References
Seasons
Species
SpeciesActivities
SpeciesLifestage

List of forms:

Name
frmActivities
frmAssociations
frmChart
frmEcoRegions
frmGrids
frmHabitats
frmInfluences
frmLifestages
frmMain
frmOccurence
frmOtherActivities
frmPlaceTime
frmPlans
frmPredators
frmPrey
frmReferenceInstance
frmReferences
frmSeasons
frmSpecies
frmSpeciesActivities
frmSpeciesLifestage
frmTestPlot
sfLevel1Habitats
sfLevel2Habitats
sfLevel3Habitats

APPENDIX 11C. THE DATABASE DESIGN PRINCIPLES

One of the primary aims of relational database design is to provide a system that is based around real physical entities and processes. If this principle is adhered to, it is much easier to develop a database system that is understandable to users and maintains data integrity. It also allows for much greater flexibility in analyses and future alterations and additions to the system. A critical aspect is that the complexity of the natural system being analyzed can be represented in terms of the data content rather than the data structures. Providing this is achieved then a deceptively simple system can be a powerful tool for both the environmental researcher and manager alike. It means that the resources used to both collect the data, and design the system to manage it, have been put to the best possible use. It also allows for the more effective integration with companion systems.

The integrity of the relational database is maintained through an extensive number of primary and foreign keys. The primary keys prevent the illogical addition of duplicate records. Though obviously sensible in itself, this becomes particularly important at the analysis stage since such duplicate values can cause multiplication of query results. Correctly normalized tables (to third normal form) and foreign keys that prevent many-to-many relationships between tables also guard against such errors in analysis.

Enforcing referential integrity via foreign keys also ensures the correct grouping of results during stratified analyses. These safeguards enforce certain requirements at the data entry stage. Basically these boil down to always first having a correct reference value in the reference tables before such a value can be used in the main data entry tables. For example, you cannot enter a species name in the SpeciesActivities section unless it first exists in the Species table. The same principle applies to life stages, habitat levels, grids and eco-regions, management plans and seasons and other activities. Even if one of these entities is not being used in a particular data element, at least one value such as ‘All’, ‘Unknown’ or ’Not-applicable’ must be entered in the relevant table. The system will not let you proceed with routine data entry until you have done this.

The values in these reference tables thus ensure the values entered during routine entry of the mass of data are consistent and correct. The reference values are also the source of choices offered in the drop down combo boxes which offer a choice of values to enter at both the table and form level. This saves on having to remember and type values correctly.

Having the data values presented in this way also means that full descriptive terms can be used instead of having to use meaningless codes and abbreviations. This both simplifies the database design and makes the system clearer to all users.

There are also simple rules enforced governing the allowable values for various attributes. Generally these allow either null values or ones that are within applicable physical ranges.

A system based around a sound fundamental data model is far simpler and thus comprehensible even to the non-database specialist. It also makes the definition of analyses far simpler; negating

the need for hidden code modules. This gives the user far greater scope to use the system themselves as a research and management tool without constant recourse to a computing specialist.

If data are to be entered at different sites, then careful planning must be made as how to coordinate these sites to ensure the resultant data sets can be combined without compromising data integrity. The simplest option is to enter all the data into one database. It can be set up for multiple users to do this. The users can connect to it either via a local or wide area network or via the internet. For the latter option it would be necessary to develop the 'Active Server Pages' that would be required as an interface for internet data entry. The other possibility is for the database to be 'replicated' and later 'synchronized.' These strategic decisions need to be taken, communicated and enforced by those responsible for managing the database and adhered to by those using it!

APPENDIX 11D. EXAMPLE METHODOLOGY FOR GENERATING SPATIAL AND TEMPORAL DATA FROM SOURCE DOCUMENTS.

This example methodology is intended to demonstrate how spatial and temporal patterns could be extracted from the 'Updated Life History Descriptions.doc' document, if as and when this were required, and represented as hard data in the Habitat Use Database, that would then have the capability of being analyzed. For the time being these methods are not required because the database concentrates simply on mapping habitats that are capable of being matched to GIS substrates. It is, however, worth reading these sections since the principles explained are also applicable to most of the other attributes in the database, and how they all fit together in the overall framework.

The researcher should first decide on definite scales of spatial and temporal sub-division, e.g. four seasons and suitable grid squares. Then for each individual species using a chart of the West Coast region with these grids marked and isobaths marked proceed to mark on the stated ranges (maximum and preferred:- note: an additional range association field would be needed to reflect this). Also from the 'Habitat' sub heading in the document mark on the depth preferences within the range, what life stage they are, what season it is, and what they are doing at that time. Additional information on this score should also be gleaned from the sections on 'Movements and Migrations' and 'Reproduction' sections of the document.

Those plots should then be used as the basis for building up the bank of descriptive records. This should be done grid square by grid square and season by season within each grid square.

Thus wherever there is a grid square where the species occurs, we create the first record for the species. This record will list the grid square ID, the season (or value for whatever temporal attribute you have agreed upon). It need not list or assume particular values for the four habitat fields unless these are explicitly known, because this information will probably be sourced from the GIS info. However, where definite habitat data are available, they should be entered as they could later be used to refine the distribution within the grid square when matched against the substrate data from the GIS system. Where multiple habitat definitions exist within the same Grid square, then multiple PlaceTime records should be created to represent this.

All the other relevant data that are available for this grid square, at that time of year, should also be entered, i.e. any Place / Area name, EcoRegion, Lat-Long and possibly year. The depth temp salinity and oxygen values should again be gleaned of seasonal oceanographic charts where possible.

Anything can be used as a PlaceTimeID providing it is a unique value. Previous extensive discussion has agreed that this should be composed of a complex code combining the values from each of the attributes. Though such a code is never processed during analysis it is useful for comprehension during data entry and review.

e.g. for gridsquare, season, hab1, hab2, hab3, hab4, we could have a code such as G15_Sa_H1c_H2b_H3h_H4b or some such.

*(Data on 'Influences,' e.g. fishing activity, could be entered as well should you choose to use this feature. If so, duration would have to be summed according to the temporal scale that is being used, e.g. average days of fishing in the season per unit of fishing gears * 'average' numbers of fishing gears operating in that grid square during that season.)*

Then drop down into the 'Occurrence' sub form and enter the Species, Gender and Lifestage for that particular instance only.

Under 'SpeciesActivity' list the type of activity for that Species-Gender-Lifestage and likewise enter any details concerning predation and prey from the section on 'Trophic Interactions.

Enter additional records in this section for any other genders and lifestages that occur in that grid square in that season for THAT species.

Don't bother with the details for any other species at this stage as each species will be done in turn.

Then move on to the next 'PlaceTime' definition. This could be the same grid square and season but a different combination of habitat definitions within these or it might be a new season within the grid square or a new grid square altogether.

Repeat the whole process building up the description of the system Species by species, grid square by grid square, season by season, habitat by habitat, gender by gender, life stage by life stage, activity by activity.

Note that the easiest way to do this is by using the PlaceTime Centric form even though we are progressing species by species from the 'Updated Life Histories' document. Obviously as PlaceTime(habitat) definitions are built up these can be reused where applicable for other species and can be retrieved via the code and/or order of sorting provided in the drop down menu choices.

The following ‘scenarios’ will, hopefully, help explain how this method of data ‘extraction’ enables increasing complexity in the natural system to be encapsulated as an increase in data rather than an increase in data structures and database complexity. The principles are equally applicable when designing a survey to gather primary source data as they are here for use in ‘extracting’ data from secondary descriptive material.

Any given situation from the very simplest to the most complex is represented within identical data structures. The only difference is the <u>amount</u> of data required to describe the situation.

In the simplest case the entire environment could be described by a single record. There would be one life stage for one species occupying a uniform space for all time. If we introduced a

second life stage, that would double the number of records. If we then introduce a second species, also with two life stages, that would double the number of records again.
If we divide the area up into five eco-regions then that potentially increases the data by 5 times (not allowing for variable spatial distribution).
If we introduce two habitat types, that again would double the number of records (where both habitats occur).
If we introduce a 'year' then the data set is multiplied for each year recorded (not allowing for variable temporal distribution).
If we introduce a season then the number of records required is multiplied by the number of seasons (again not allowing for seasonal patterns) and so on for each new attribute that we introduce.
Each of these increases in complexity requires no alteration whatsoever to the structure of the database.

The same kind of principle applies to the linked subsidiary tables describing species activities, predators and prey.

It is useful to bear that 'scenario' in mind when breaking down the descriptive 'Updated Life Histories' document into data that is capable of analysis with this system.

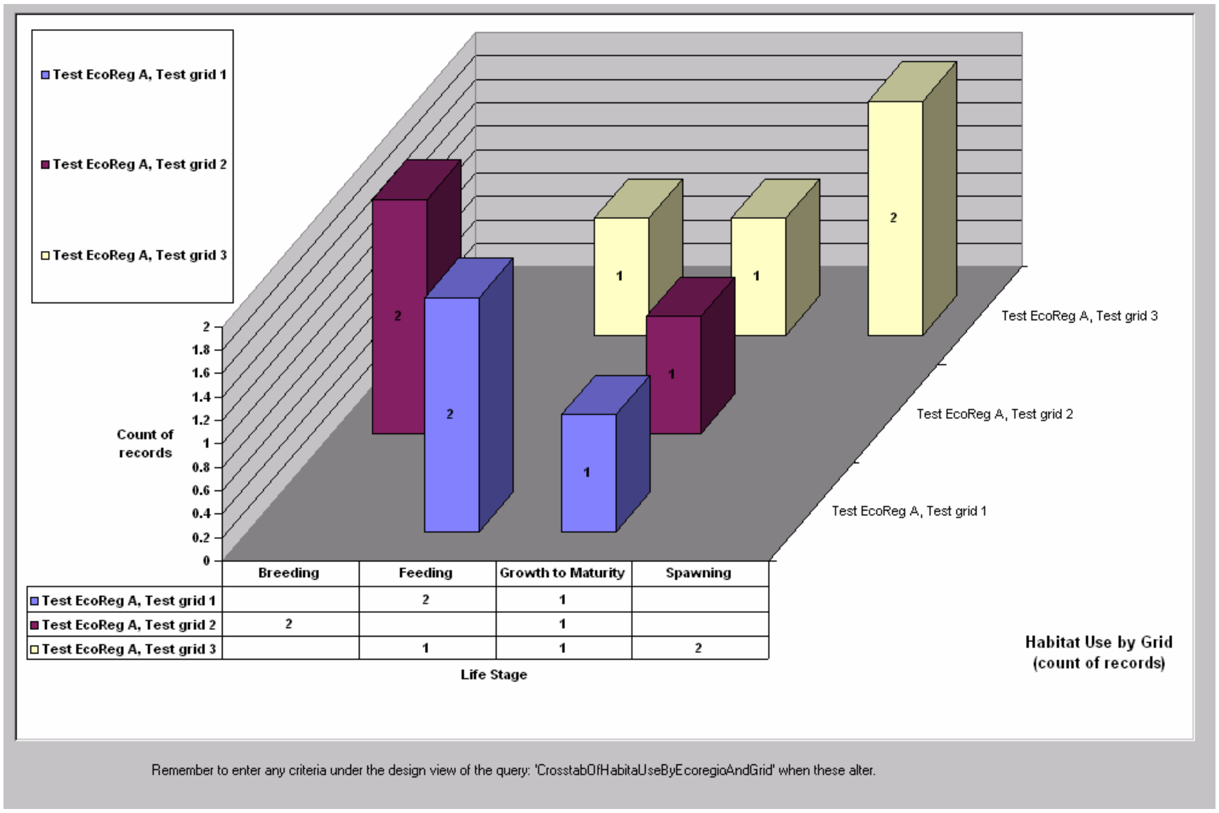
Thus, if it is intended to, say, break down analyses by EcoRegions, then these must be looked for in the information available. Even if a given Species-Lifestage genuinely occupies a given habitat throughout the entire West Coast, five records must be entered to describe it correctly; one per eco-region. That would mean in practice there being five occurrence records being entered for the SpeciesLifestage each with a different PlaceTimeID. Those related PlaceTime records would be identical apart from having

a) a different value under the EcoRegion field, and
b) a different PlaceTimeID code.

Of course in reality it is more likely that the SpeciesLifestage may for example only be recorded in three of the five EcoRegions. In this fashion real complex patterns of distribution can be correctly represented.

The principles outlined above for EcoRegions are equally applicable when dealing with Grids, Seasons, Years and the various combinations of habitats.

Here is an example of the charted output from a query analyzing test data for spatial distribution of species activity across a grid scheme within an Eco-Region for a particular species.



If only habitat variations are intended to be used for analyses then obviously that reduces the amount of data required, there not being the need to break things down into their EcoRegion and Seasonal components.

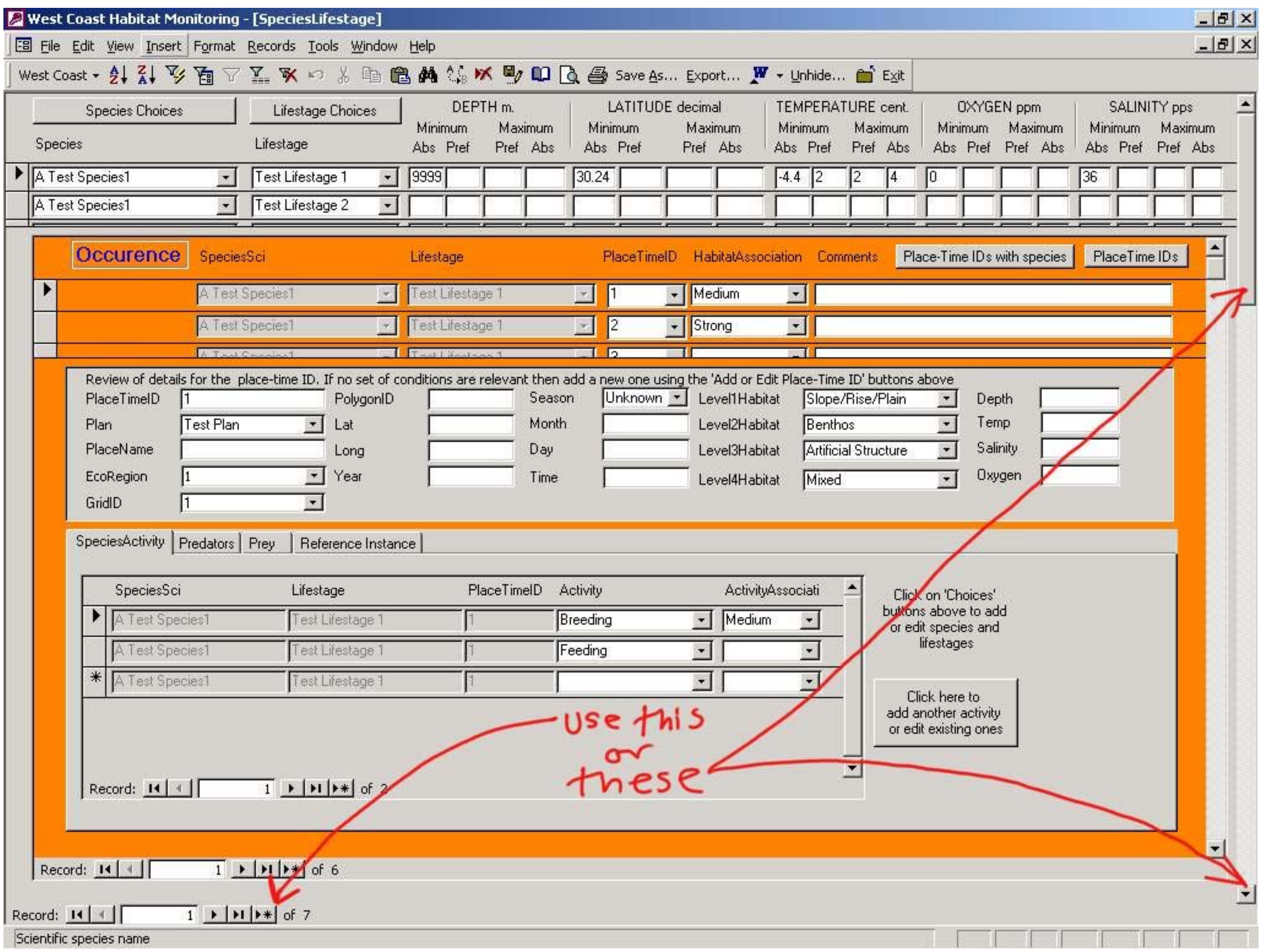
APPENDIX 11E. TUTORIAL FOR EXAMPLE DATA ENTRY

There follows a short tutorial of how data were extracted from the Updated Life Histories document for Petrale sole where it was confined to substrate classification, latitude, depth, salinity and temperature ranges. Temporal and spatial variation was ignored for the present.

The species names should all be in there to start with but you would in theory first go to species and check name. Use the Binocular 'find' symbol on the tool bar to search for the name you are looking for. Make sure the 'Look In' and 'Match' options are set correctly. The scientific names are also in alphabetical order in any case.

From the opening form 'frmMain' open the 'Data: Species centric' form by selecting that button.

Chose the new record button from the navigation buttons of the 'outer', 'parent' Species-Lifestage form as illustrated below:



Then click the drop down box for the 'Species Box' and pick out the species name for entry: Eopsetta jordani in this case.

Decide whether you are going to enter the Species Lifestage attributes to represent all life stages, a selection of life stages, or all the life stages for which information is available. According to your choice you will have one record or a number to enter (one per life stage chosen). Go through the document trawling out the values for the four range limits for depth, latitude, temperature, oxygen and salinity. This is probably best done by using the word search facility for the key word in each case for the species under consideration.

For Petrale Sole the initial depth information under the 'Fishery' and 'Habitat' sections indicates that adults have a preferred range of 300 to 460m but have an absolute range of 0 to 550m. The fields are filled in accordingly. A new record is created for the juvenile life stage. The details for each of the physical characteristics can be edited in for each of the life stages at the same time or each life stage can be completed separately for all of the characteristics needed for each field before moving onto the next life stage. Whichever is most convenient for the data enterer depending on the order the data is extracted from the descriptive document.

Remember entire records can be copied and pasted into the next row as a new record in order to save retyping. You obviously have to then edit the necessary key fields (e.g. here this would most likely be the 'Lifestage' field) so that the record is not a duplicate before that new record can be saved. It goes without saying that you would also amend any of the data in the fields for the physical characteristics where these were different from the previous record. The field above can also be copied where this is simpler by simply holding down the Ctrl and 'C' keys simultaneously in order to save retyping or selecting from a drop down list.